THIS WEEK

EDITORIALS

WORLD VIEW Reform PhD programmes or close them down **p.261**

SINUS VIRUS Lab culture breakthrough could offer cold comfort **p.262**

PENGUINS Populations plummet as Antarctic ice melts **p.263**

Accidents in waiting

Every researcher and institution should question their own attitudes to safety in the lab after the death of an undergraduate student in a Yale University workshop.

The shocking death of physics and astronomy undergraduate student Michele Dufault in a machine shop at Yale University in New Haven, Connecticut, last week should grab the attention of researchers and safety officers at universities across the United States, and the wider world. Rightly, the immediate focus is on whether the university could have taken more precautions to prevent the accident. But whatever the verdict, Dufault's death — late at night and probably while working alone — should remind every researcher to consider their own attitude to safety, and whether it is crowded out by other priorities.

Most scientists are well aware of poor safety practices in their laboratories — such as too many people working on their own, students not properly trained to use equipment, or a general reluctance to wear safety glasses and lab coats. But, just as bottles of unidentified solvents can be stashed guiltily in the depths of a fume cupboard, so such problems are often pushed to the back of the mind, and only properly confronted after an accident.

Most worrying is that it seems researchers only change their attitudes to safety when affected directly by an accident, such as in their own laboratory. A tragedy elsewhere is not always sufficient motivation. After the 2009 death of research assistant Sheharbano Sangji following a fire in the chemistry department at the University of California, Los Angeles (UCLA), safety policies there, such as snap lab inspections to make sure researchers wear protective coats, were reviewed and tightened. But the impact outside the University of California system is hard to determine, and there is little evidence that other chemistry laboratories have responded by changing their practices (see page 270). This could soon change — the deaths of both Sangji and Dufault will feed into ongoing federal-level inquiries into laboratory safety by the US Chemical Safety and Hazard Investigation Board and the National Research Council.

It is certainly true that laboratory researchers work in much safer conditions now than during the more reckless days of the 1950s and 60s. But such improvements must not breed complacency. For years, environmental health and safety officers have complained that there is no good source of consistent data on laboratory accidents, which could be studied to determine effective safety interventions. As such, it is impossible to say with certainty which laboratories perform well on safety and which badly. And no one can rigorously compare academia's accident rate to that of industry.

UCLA has just announced a centre to study laboratory safety, which may start to pull together some of this information. That would be a good start, and many of the data it would need for the task already exist. Individual research departments in both academia and industry often keep statistics far more detailed than required by federal authorities — recording minor incidents and near misses, as well as major accidents. The American Chemical Society's health and safety division has started informally tracking statistics of deaths in academia and industry, to see if any discernible patterns arise. Last year, it surveyed the safety culture in chemistry laboratories, and it plans to repeat the exercise.

The UCLA centre would be a good place to pool this knowledge and make it widely available — and not just between health and safety officers, who already discuss experiences and data. There are useful examples of collaboration on safety issues elsewhere, such as an MIT peer-review process with the National University of Singapore, in which each institution audited the health and safety programmes of the other. If asked, researchers are usually more than happy to make safety policies available or send them out to others. Imperial College London and UCLA both make their safety plans widely available and grant requests for information from researchers in other countries.

> **"Poor safety practices are often pushed to the back of the mind and only confronted after an accident."**

To see safety precautions as a drag on research is an irresponsible and counterproductive attitude, but one that is hard to change. At UCLA, for example, too many researchers see newly introduced safety officers as 'police' to skirt round, rather than experts with whom to collaborate.

Leaders of research projects must take responsibility for the safety of the scientists doing the work, and must start to work with safety officers, rather than endure them. In turn, senior figures in academic departments must realize that practices and priorities have changed since their earlier days, and be willing to shut down laboratories until any potentially dangerous working practices are improved.

The circumstances that ended the life of Michele Dufault last week may have been unusual, even unique. But universities and researchers who feel that there are no lessons to learn from such accidents are a danger to themselves and others. ■

Fix the PhD

No longer a guaranteed ticket to an academic career, the PhD system needs a serious rethink.

The world has many problems and it will take a lot of bright, educated people to solve them. So, on the face of it, it seems like a good thing that more and more people are earning PhDs in science, technology and engineering. Most countries, convinced that higher education and scientific research are key to economic growth and prosperity, are expanding doctoral education in science. The thought, as one researcher who has studied doctoral-education trends puts it, is that you can "grow PhDs like mushrooms".

The consequence of that mushrooming depends on where it is taking place, and in which discipline, as our overview of PhD systems

around the world shows (see page 276). Clearly, such expansion results in an extraordinary amount of good research (see page 283). And in the rapidly growing tiger economies, for example, most of those with PhDs quickly find good jobs.

But there are reasons for caution. Unlimited growth could dilute the quality of PhDs by pulling less-able individuals into the system. And casual chats with biomedical researchers in the United States or Japan suggest a gloomy picture. Exceptionally bright science PhD holders from elite academic institutions are slogging through five or ten years of poorly paid postdoctoral studies, slowly becoming disillusioned by the ruthless and often fruitless fight for a permanent academic position. That is because increased government research funding from the US National Institutes of Health and Japan's science and education ministry has driven expansion of doctoral and postdoctoral education — without giving enough thought to how the labour market will accommodate those who emerge. The system is driven by the supply of research funding, not the demand of the job market.

The problem is widely discussed, yet many PhD programmes remain firmly in the traditional mould — offering an apprenticeship for academic research, even as numbers of academic positions stagnate or decline. Yes, there are many worthwhile careers outside academia for science PhD holders (*Nature* would be down to a skeleton staff without them). And most people with science PhDs eventually find satisfying jobs. But they probably feel that spending years performing minipreps was not the most appropriate way to become a banker or a teacher. Widening concerns about dismal job prospects are dissuading some of the brightest candidates from taking the PhD route.

Something needs to change — but what? Ideally, the system would produce high-quality PhD holders well matched to the attractive careers on offer. Yet many academics are reluctant to rock the boat as long as they are rewarded with grants (which pay for cheap PhD students) and publications (produced by their cheap PhD students). So are universities, which often receive government subsidies to fill their PhD spots.

One way in which governments can bring about change is to better match educational supply with occupational demand. They should get smart, independent labour economists to comb through wage and employment data that reveal which types of science-related job are in short supply, and talk to stakeholders on the ground to confirm the findings. Governments should then open the doors to more PhDs only where they are most needed. Such analyses are already under way, and should be encouraged.

A second route is to reform the PhD itself (see page 261), and reset the expectations of those in the system. Imagine bright young things entering a new kind of science PhD, in which both they and their supervisors embrace from the start the idea that graduates will go on to an array of demanding careers — government, business, non-profit and education — and work towards that goal (see page 381). The students meet supervisors from a range of disciplines; they acquire management, communication, leadership and other transferable skills alongside traditional academic development of critical thinking and analysis; and they spend six months to a year abroad.

"Widening concerns about dismal job prospects are dissuading the brightest candidates from the PhD route."

Some such efforts have already begun: for example, US institutions vie to win prestigious grants from the Integrative Graduate Education and Research Traineeship (IGERT) programme run by the National Science Foundation, which promotes highly interdisciplinary PhDs (see page 280)

The IGERT scheme shows how appropriate reward structures can drive change. Governments and funding agencies should require educational institutions to release figures showing how many of their PhD students complete the course, and how many go on to find employment and where, and should award some proportion of funding accordingly. This would also help prospective students to select a good course, and force worse-performing programmes to shape up or close.

Until any of this becomes commonplace, it is up to prospective graduate students to enter a science PhD with their eyes open to the opportunities — or lack of them — at the end. Not all mushrooms grow best in the dark. ■

Home truths

A new report offers useful insight into the continuing stalemate over global warming.

In just over six months' time, officials from the world's nations will meet under the auspices of the United Nations to try again to complete the task that was beyond them in Copenhagen in 2009, to establish a legally binding treaty to curb global warming. It is hard to see why it could go any better this time — if anything, the global economic slump and the failure to pass cap-and-trade legislation in the United States will make it even harder. A report published this week in the United States does an excellent job of probing the reasons for this stalemate, and shines light on some uncomfortable truths. It should be essential reading for anyone with a passing interest in the climate-change debate.

The report, *Climate Shift: Clear Vision for the Next Decade of Public Debate*, is written by Matthew Nisbet, a professor of communication and environmental science at American University in Washington DC. It focuses on the situation in the United States, and particularly its political failure to pass comprehensive climate legislation. But the points it makes go far beyond Capitol Hill. And it effectively dismantles three of the most common reasons given by those who have tried, and failed, to garner widespread support for policies to restrict greenhouse gases.

First — the failure of the US Senate to pass a cap-and-trade bill in 2010 cannot be blamed directly on the financial lobbying muscle of the conservative movement and its allies in industry. In 2009, the report says, although a network of prominent opponents of cap and trade, including ExxonMobil and Koch Industries, spent a total of US$272 million lobbying policy-makers, environmental groups in favour of cap and trade mobilized $229 million from companies such as General Electric and other supporters to lobby for environmental issues. Indeed, the effort to pass cap and trade, Nisbet notes, "may have been the best-financed political cause in American history".

Second — most of the mainstream media coverage of climate change gets it right. During 2009 and 2010, Nisbet writes, around nine out of ten news and opinion articles in *The New York Times*, *The Washington Post* and CNN's online site reflected the consensus scientific position. *The Wall Street Journal* regularly presented the opposite view in its opinion pages, but eight out of ten news items still backed the science.

Third — conservative media outlets such as Fox News and controversies such as the coverage of e-mails hacked from the University of East Anglia in the United Kingdom have a minimal impact on public attitudes to climate change, because such influences tend to only reinforce the views of those who already hold doubts.

The failure of cap and trade in the United States, Nisbet concludes, was not down to poor communication, but was due to framing the issue of greenhouse-gas emissions as a problem that could be solved by a specific policy. More useful, he says, would be to present climate change as an issue that needs to be addressed at many levels, similar to public health or poverty. Those, of course, are far from ideal models — but we live in far from ideal times. ■

➲ NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv

WORLD VIEW *A personal take on events*


R. HOWARD

Reform the PhD system or close it down

There are too many doctoral programmes, producing too many PhDs for the job market. Shut some and change the rest, says **Mark C. Taylor**.

The system of PhD education in the United States and many other countries is broken and unsustainable, and needs to be reconceived. In many fields, it creates only a cruel fantasy of future employment that promotes the self-interest of faculty members at the expense of students. The reality is that there are very few jobs for people who might have spent up to 12 years on their degrees.

Most doctoral-education programmes conform to a model defined in European universities during the Middle Ages, in which education is a process of cloning that trains students to do what their mentors do. The clones now vastly outnumber their mentors. The academic job market collapsed in the 1970s, yet universities have not adjusted their admissions policies, because they need graduate students to work in laboratories and as teaching assistants. But once those students finish their education, there are no academic jobs for them.

Universities face growing financial challenges. Most in the United States, for example, have not recovered from losses incurred on investments during the financial fiasco of 2008, and they probably never will. State and federal support is also collapsing, so institutions cannot afford to support as many programmes. There could be an upside to these unfortunate developments: growing competition for dwindling public and private resources might force universities to change their approach to PhD education, even if they do not want to.

MOST DOCTORAL PROGRAMMES CONFORM TO A MODEL DEFINED IN THE MIDDLE AGES.

There are two responsible courses of action: either radically reform doctoral programmes or shut them down.

The necessary changes are both curricular and institutional. One reason that many doctoral programmes do not adequately serve students is that they are overly specialized, with curricula fragmented and increasingly irrelevant to the world beyond academia. Expertise, of course, is essential to the advancement of knowledge and to society. But in far too many cases, specialization has led to areas of research so narrow that they are of interest only to other people working in the same fields, subfields or sub-subfields. Many researchers struggle to talk to colleagues in the same department, and communication across departments and disciplines can be impossible.

If doctoral education is to remain viable in the twenty-first century, universities must tear down the walls that separate fields, and establish programmes that nourish cross-disciplinary investigation and communication. They must design curricula that focus on solving practical problems, such as providing clean water to a growing population. Unfortunately, significant change is unlikely to come from faculty members, who all too often remain committed to traditional approaches. Students, administrators, trustees and even people from the public and private sectors must create pressure for reform. It is important to realize that problems will never be solved as long as each institution continues to act independently. The difficulties are systemic and must be addressed comprehensively and cooperatively. Prestige is measured both within and beyond institutions by the number and purported strength of a department's doctoral programmes, so, seeking competitive advantage and financial gain from alliances with the private sector, universities continue to create them. As is detailed on page 276, that has led most fields to produce too many PhDs for too long.

➲ **NATURE.COM**
Comment on the future of the PhD:
nature.com/phdfuture

The solution is to eliminate programmes that are inadequate or redundant. The difficult decisions should be made by administrators, in consultation with faculty members at their own and other universities, as well as interested, informed and responsible representatives beyond the academic community who have a vested interest in effective doctoral education. To facilitate change, universities should move away from excessive competition fuelled by pernicious rating systems, and develop structures and procedures that foster cooperation. This would enable them to share faculty members, students and resources, and to efficiently increase educational opportunities. Institutions wouldn't need a department in every field, and could outsource some subjects. Teleconferencing and the Internet mean that cooperation is no longer limited by physical proximity.

Consortia could contain a core faculty drawn from the home department, and a rotating group of faculty members from other institutions. This would reduce both the number of graduate programmes and the number of faculty members. Students would have access to more academic staff with more diverse expertise in a wider range of fields and subfields. Faculty members will resist, but financial realities make a reduced number of posts inevitable.

Higher education in the United States has long been the envy of the world, but that is changing. The technologies that have transformed financial markets and the publishing, news and entertainment industries are now disrupting the education system. In the coming years, growing global competition for the multibillion-dollar education market will increase the pressure on US universities, just when public and private funding is decreasing. Although significant change is necessary at every level of higher education, it must start at the top, with total reform of PhD programmes in almost every field. The future of our children, our country and, indeed, the world depends on how well we meet this challenge. ■

Mark C. Taylor *is chair of the department of religion at Columbia University in New York and the author of* Crisis on Campus: A Bold Plan for Reforming Our Colleges and Universities *(Knopf, 2010).*
e-mail: mct22@columbia.edu

RESEARCH HIGHLIGHTS

Selections from the scientific literature

ASTRONOMY

Sharper cosmic pencils

In order to look back in time, astronomers peer deep into a narrow 'pencil beam' of space. These narrow surveys can be easily confounded by variations in the density of matter on scales larger than the pencil beam.

Benjamin Moster, then at the Max Planck Institute for Astronomy in Heidelberg, Germany, and his colleagues have developed a recipe for measuring uncertainty in density. The team combined measurements of galaxies with models of dark matter, a mysterious material thought to bind galaxies together, and was able to predict how the density variation affects galaxy measurements at different distances from Earth. The results should ensure more accurate interpretation of past and future pencil-beam surveys.

Astrophys. J. 731, **113 (2011)**

VIROLOGY

Stubborn virus grows in sinus

Five years after its discovery, human rhinovirus C, a cold-causing virus that is associated with acute respiratory illness in children, has finally been cultured. This is the first step towards finding the virus's infection mechanisms, as well as possible antiviral targets.

Yury Bochkov of the University of Wisconsin–Madison and his colleagues grew two forms of HRV-C from nasal lavage fluid on sinus tissue left over from human surgeries. They then cloned a copy of the viral genome in a bacterial DNA structure called a plasmid, from which they produced new viral RNA. After introducing this construct into standard cell lines, they recovered infectious virus particles.

HRV-C seems to attach to cells by way of an unknown receptor not used by other HRVs, information which could be used to guide research into treatment approaches.

Nature Med. **doi:10.1038/nm.2358 (2011)**

M. MULDER/R. NETHERLANDS INST. SEA RES.

ECOLOGY

Jonah and the waterfowl

Tiny sea snails can survive a voyage through a duck's gizzard — where prey are typically crushed to death — only a little the worse for wear.

Gerhard Cadée at the Royal Netherlands Institute for Sea Research in Texel sifted through the faeces of shelducks, *Tadorna tadorna*, for shells of the snail *Hydrobia ulvae* **(pictured)**. Snails that suffered only minor damage quickly repaired their gashes after Cadée placed them in a laboratory aquarium. The creatures developed scars on their shells that resembled marks left by certain predators that had tried and failed to eat them.

Between 2.8% and 41.8% of the *H. ulvae* shells collected at 11 sites along the Wadden Sea bore repair scars. Such wide variation may complicate efforts to relate the frequency of scars in fossilized shells to levels of predation, Cadée says.

Palaios 26, **245–249 (2011)**

STEM CELLS

New ways to change fates

Induced pluripotent stem cells (iPSCs) hold enormous promise for regenerative medicine: they have the potential to develop into cells of any type, just as embryonic stem cells do, but do not require the controversial use of human embryos. Until now, the most efficient way to make iPSCs from human skin cells required a virus to ferry four 'reprogramming' factors into cells.

Edward Morrisey at the University of Pennsylvania in Philadelphia and his colleagues describe a recipe for making iPSCs that does not require these four factors, and boosts the efficiency of the reprogramming process. By infusing cells with a viral vector encoding certain small, RNA strands called microRNAs, the team was able to reprogram human skin cells two orders of magnitude more efficiently than is possible with the standard method.

Cell Stem Cell 8, **376–388 (2011)**

OPTICAL PHYSICS

A liquid of photons

For decades, physicists have mused over the theoretical properties of Luttinger liquids, strange one-dimensional quantum states in which charge and magnetic spin of particles are predicted to decouple and move at different speeds.

Dimitris Angelakis at the Technical University of Crete, Greece, and his colleagues have taken a step towards creating a Luttinger liquid in the lab. They propose a practical scheme whereby two different types of atom would be trapped in a hollow, one-dimensional optic fibre and tuned with two sets of opposing laser beams. When one pair of beams is switched off, the authors show, a photon pulse in which charge and spin are decoupled will move along the fibre.

Phys. Rev. Lett. 106, **153601 (2011)**

COMMUNITY CHOICE
The most viewed papers in science

NEUROSCIENCE

Sleep deprivation: a risky business

HIGHLY READ on jneurosci.org in March

What do stock traders, doctors and soldiers have in common? Each often needs to make important decisions on little sleep.

Michael Chee of the Duke-NUS Graduate Medical School in Singapore and his colleagues used brain imaging to study how sleep deprivation affects decision-making. They scanned the brains of healthy adults performing a complicated gambling task. When the volunteers had stayed up all night, their strategy shifted towards seeking bigger gains over protecting against losses. The shift in behaviour correlated with changes in brain activity in regions previously associated with reward evaluation and regret.

Interestingly, the shift in strategy did not correlate with vigilance, suggesting that although stimulants such as caffeine might make sleepy people more alert, they do not improve decision-making abilities.

J. Neurosci. 31, **3712–3718 (2011)**

S. RENTMEESTER/FMAX

MECHANICS

A stable, riderless bicycle

A riderless bicycle can travel for some distance without falling — an ability generally assumed to be due to the gyroscopic forces of the spinning front wheel keeping it upright. But this is not the whole story, according to a paper by Andy Ruina of Cornell University in Ithaca, New York, and his colleagues. Following a theoretical model, the team built a bicycle (**pictured**) in which the gyroscopic forces were cancelled by a second set of wheels, off the ground, that spun in the reverse direction.

The authors found that the bike was relatively stable at speeds of more than 2 metres per second, showing that gyroscopic forces are not necessary for bike stability. In the case of their custom bike, the low steering assembly — handlebars, fork and front wheel — fell faster than the rest of the bike, tending to steer the bike in the direction of the fall, just as a rider would do.

Science 332, **339–342 (2011)**

BEHAVIOUR

Diversity beats ability

Can you guess how many marbles are in a jar? Maybe you are a savant or maybe not; either way, odds are that, between them, a group of people will beat your guess.

Stefan Krause at the Lübeck University of Applied Sciences in Germany and his colleagues asked visitors to a museum to play the marble guessing game, and collected more than 2,000 guesses. When they made random groups out of these results, they found that the average guess of groups with 40 or more members was better than the best quarter of individual guesses, which implies that large groups of average ability are smarter than individual stars. The result suggests that, like schooling fish or swarming bees, humans can make fruitful use of collective decision-making, and that groups might outperform individuals of high ability — hinting at a selection pressure for diverse personalities.

Anim. Behav. 81, **941–948 (2011)**

CANCER BIOLOGY

Laying siege to cancer

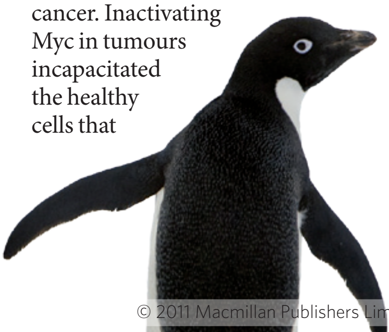
Tumours can be shrunk by inhibiting a naturally occurring transcription factor called Myc to make the tumour microenvironment inhospitable to invasive cells.

Gerard Evan, now at the University of Cambridge, UK, and his team tested the effects of Myc suppression in a mouse model of pancreatic cancer. Inactivating Myc in tumours incapacitated the healthy cells that support cancer cells. This resulted in collapsed blood vessels and reduced inflammatory responses, leading to oxygen starvation and death for tumour cells.

The fact that tumour cell death occurred after microenvironment collapse, even if Myc was inhibited only in cancer cells, suggests that the transcription factor functions as a communicator between tumours and their surroundings. And, as an essential contributor to tumour growth and maintenance, Myc is an attractive therapeutic target — especially because most cells can't wiggle around Myc suppression.

Genes Dev. **doi:10.1101/gad.2038411 (2011)**

S. ESZTERHAS/MINDEN PICTURES/FLPA



CLIMATE CHANGE

Sea ice penguin theory sinks

As sea ice has declined off the West Antarctic Peninsula, scientists expected to see declines in ice-dwelling Adélie penguins (**pictured left**) and increases in chinstrap penguins (**right**), which forage in ice-free waters. However, Wayne Trivelpiece and his colleagues at the US National Marine Fisheries Service in La Jolla, California, say there is now "overwhelming evidence" that both populations are declining in the region.

The authors' data, which cover populations at two sites over 30 years, suggest that sea-ice losses resulting from climate change have reduced the availability of Antarctic krill, the prey of both birds. If temperatures rise in future, sea ice and krill will decline further, and both species of penguin are likely to follow.

Proc. Natl Acad. Sci. USA **doi:10.1073/pnas.1016560108 (2011)**

MOMATIUK & EASTCOTT/CORBIS

NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

POLICY

Wolf delisted

The grey wolf will be removed from the US government's endangered species list in some northwestern states as a result of a policy initiative tagged on to the US federal budget bill, which was approved last week. Grey wolf populations have recovered significantly in northwestern states, but environmental campaigners such as the Center for Biological Diversity in Tucson, Arizona, bemoaned the fact that politicians had lifted protection rather than waiting for due process under the Endangered Species Act. For details of the budget's settlement for science, see page 267.

Israel to join CERN

Israel is set to become the first non-European member of CERN, Europe's high-energy physics research centre near Geneva, Switzerland. On 17 April, Israel's cabinet voted to join the lab. Full membership has historically been limited to European nations, but last June, CERN's council opened the door to outsiders. The council is expected to approve Israel's membership in an upcoming meeting. Brazil, Cyprus, Serbia, Slovenia and Turkey are also pursuing full membership.

Hormones in sport

Female athletes may not be eligible to compete as women if blood tests show they have natural testosterone levels in the male range, according to rules accepted by the International Association of Athletics Federations on 12 April. The decision on hyperandrogenism — in which the body produces higher than normal levels of androgen hormones, particularly testosterone — has been broadly welcomed by experts. See go.nature.com/xc5cnm for more.

AP PHOTO/H. KOMAE

Clean-up visions for Fukushima

As workers continue to douse stricken reactors at the Fukushima Daiichi nuclear plant with water, the facility's owner has laid out plans for stabilizing and cleaning up the site. On 17 April, the Tokyo Electric Power Company (whose executives are pictured at a harried press conference six days earlier) put forward a plan to stabilize the plant within six to nine months. According to this, workers will continue to pump in water to cool three damaged reactors, as well as spent fuel pools, while, in parallel, developing techniques to store and decontaminate used water on the site. The company also plans to cover the damaged reactors with temporary structures in order to limit the release of radioactivity.

European networks

The European Institute of Innovation and Technology (EIT) said on 14 April that it is considering eight areas — including biotechnology, smart cities and ageing — in which to fund new collaborative research networks. It already funds three such initiatives, known as Knowledge and Innovation Communities (KICs), focusing on topics including climate change. These are networks of industry and academic partnerships with EIT funding of €308.7 million (US$439 million) to 2013. Any new plans will feed into the European Commission's own proposals for the EIT's future, which are due in December. On 14 April the commission launched a public consultation on the subject.

Shuttles at rest

The four remaining vehicles of the US Space Shuttle fleet were assigned their final resting places on 12 April. *Atlantis* will remain at the Kennedy Space Center in Merritt Island, Florida; *Endeavour* will head to the California Science Center in Los Angeles; and *Discovery* will go to the Steven F. Udvar-Hazy Center in Chantilly, Virginia, which is part of the National Air and Space Museum in Washington DC. A flight test vehicle, *Enterprise*, will travel to the Intrepid Sea, Air & Space Museum in New York City. See go.nature.com/tiilci for more.

Virus sharing

In the event of a future flu pandemic, member states of the World Health Organization (WHO) will send samples of flu virus to laboratories and drug makers around the world, in return

for greater access to any vaccines created. The deal, announced by the WHO on 17 April, heads off the prospect of countries refusing to share samples with WHO laboratories in protest at not benefiting from resulting research patents or vaccines — as Indonesia did in 2007.

PEOPLE

Research fraud

Prosecutors in the United States are seeking to extradite a Danish scientist researching the relationship between autism and vaccines, who, they allege, stole more than US$1 million in research funding. Poul Thorsen was a visiting scientist at the Centers for Disease Control and Prevention (CDC) in Atlanta in the 1990s. US prosecutors say that after returning to Denmark in 2002, Thorsen submitted false invoices from the CDC to Aarhus University, which unknowingly transferred funds to his personal account. He was last week charged with 13 counts of wire fraud and 9 of money laundering.

New chief scientist

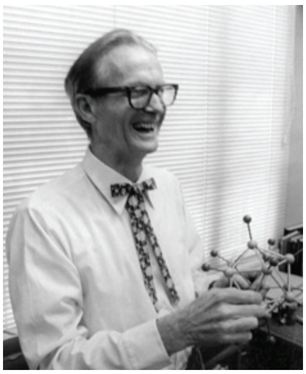
Australia's government has appointed Ian Chubb as its chief scientist. Originally a neuroscientist, Chubb has spent the past few decades in senior administration roles at various universities and research councils; most recently, he was vice-chancellor of the Australian National University in Canberra from 2001 to 2010. He replaces Penny Sackett, who in February announced her surprise resignation, halfway through her five-year term. Chubb's three-year term starts on 23 May.

Lab death

Michele Dufault, a 22-year-old undergraduate student, was found dead after an accident at Yale University's Sterling Chemistry Laboratory on 13 April. See page 270 for more.

Nobel chemist dies

William Lipscomb, who won the 1976 Nobel Prize in Chemistry for his work on



EMILIO SEGRE VISUAL ARCHIVES/AM. INST. PHYS./SPL

chemical bonding, died on 14 April aged 91. Lipscomb (**pictured**) helped to elucidate the nature of bonding between molecular clusters of boron and hydrogen atoms — called boranes — which did not obey principles known at the time. After starting out at the University of Minnesota, he moved to Harvard University in Cambridge, Massachusetts, in 1959, where he remained for the rest of his career.

RESEARCH

Brain atlas debuts

A genetic and anatomical map of the human brain, bankrolled by Microsoft co-founder Paul Allen, was officially unveiled on 12 April. The Seattle, Washington-based Allen Brain Science Institute's human brain atlas (www.brain-map.org) logged gene-expression patterns and biochemical activity at 1,000 locations in brains donated by two people, generating a total of 100 million data points. The US$55-million project follows a mouse brain atlas, released in 2006, and a map of the mouse spinal cord two years later. See go.nature.com/l9923y for more.

FUNDING

Golden rice funds

The Bill and Melinda Gates Foundation is giving US$18.6 million to research on transgenic, nutritionally fortified rice and cassava. The International Rice Research Institute in Los Baños, the Philippines, won $10.3 million to develop golden rice, which delivers extra vitamin A; it hopes that the rice will receive regulatory approval in the Philippines in 2013 and in Bangladesh in 2015. The Donald Danforth Plant Science Center in St Louis, Missouri, was given $8.3 million for work on BioCassava Plus, which contains extra vitamin A, iron and protein. The centre hopes the enhanced cassava will gain approval in Kenya and Nigeria by 2017. See go.nature.com/uuyc6o for more.

Chernobyl shelter

An international fund-raising effort to help decommission the Chernobyl nuclear power station in Ukraine seemed on 19 April to have fallen short of its goal. After a meeting in Kiev, Viktor Yanukovich, president of Ukraine, said that world governments and international organizations had pledged an extra €550 million (US$780 million) to help build a spent-fuel storage facility and an enormous steel arch to cover the shattered reactor — currently surrounded by a crumbling concrete sarcophagus. But the meeting had hoped to raise €740 million to make up the roughly €2-billion cost of this effort. See go.nature.com/xmmfhk for more.

COMING UP

24–29 APRIL
The European Science Foundation is holding a week-long conference devoted to the science and technology of graphene, in Obergurgl, Austria.
go.nature.com/xnyc4t

26 APRIL
The 25th anniversary of the nuclear disaster at Chernobyl in what is now Ukraine.

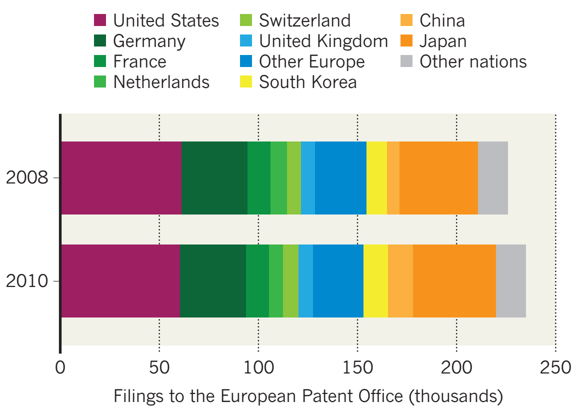
TREND WATCH

Patent filings received by the European Patent Office in 2010 topped 235,000, an all-time high. Applications from the European Union and United States slowed in the financial crisis, but have recovered, and filings from China have almost doubled from 6,490 in 2008 to 12,698 in 2010. But China's portfolio is unbalanced: 43% of its applications considered in 2005–10 were in digital and telecommunications, whereas biotechnology — a growing sector for other nations — made up only 3% of claims.

SOURCE: EPO

PATENT FILINGS RESUME GROWTH

A boost from Asia accounted for an upturn in European patent filings in 2010.



NATURE.COM
For daily news updates see:
www.nature.com/news

NEWS IN FOCUS

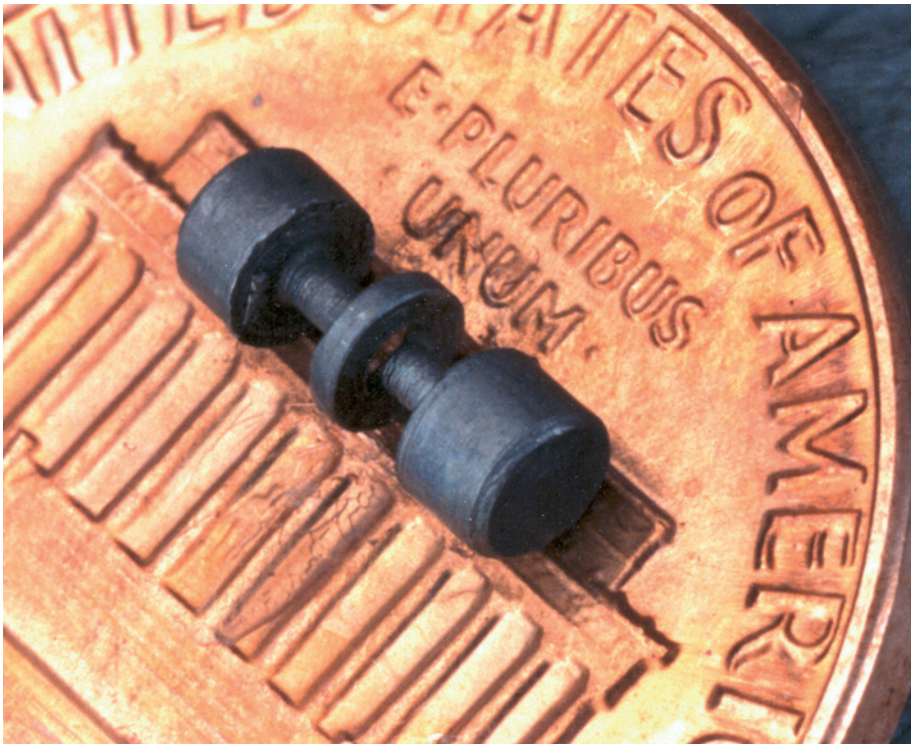
LAB SAFETY Death spurs calls for a culture change on campus **p.270**

CANCER Questioning the leading theory on metastasis **p.273**

Q&A A nuclear-safety expert gives a post-Fukushima prescription **p.274**

HIGHER EDUCATION The glut of PhDs — and what to do about it **p.276**

J. BERGQUIST, D. WINELAND/NIST



A mercury-ion trap was devised by US government scientists, who must now make do with fewer pennies.

POLICY

US budget a taste of battles to come

Deal leaves science agencies with less but avoids deep cuts.

BY EUGENIE SAMUEL REICH, IVAN SEMENIUK, JEFF TOLLEFSON AND MEREDITH WADMAN

Rush Holt calls it "deficit attention disorder". Speaking earlier this month at a science-policy luncheon on Capitol Hill, the New Jersey Democratic congressman and nuclear physicist was reflecting on the fiscal climate that has settled over Washington DC. Last week, the impact of that climate was felt acutely, as Congress passed a federal spending bill that makes the deepest cut ever to the US budget. For scientists employed or supported by government research agencies, the news would seem to be dire. The deal, which President Barack Obama signed into law on 15 April, applies the first significant funding cuts to most of those agencies in a generation — with the strong possibility of deeper cuts to come. Yet, given what might have been, the outcome is a relief to most research advocates.

"The basic-research agencies all did pretty well," says Patrick Clemins, director of the research-and-development budget and policy programme at the American Association for the Advancement of Science in Washington DC.

Back in February, the Republican-led House of Representatives proposed a budget that would have ripped billions of dollars from the science agencies. That would have left the agencies not only well below the funding levels specified by Obama when he tabled his budget request more than a year ago, but also well below the 2010 levels they have been operating under since October.

The deal restores the bulk of those lost billions. Most science agencies now face cuts of about 1% (see 'Splitting the difference'). That is a reassuringly small portion of the US$38.5 billion, or 5.8%, cut from the 2010 allocation for non-military discretionary spending from which most academic research is funded. "In the end," says Barry Toiv of the Association of American Universities in Washington DC, "it is clear they decided that while it was important to cut federal spending they would continue to prioritize research and education."

Nevertheless, the fractious debate that has enveloped Congress for the past six and a half months foreshadows a much tougher battle over the 2012 budget. Although Obama has called for investments in science and technology to "win the future", Republicans have vowed to fight for deep cuts to reduce the escalating deficit without resorting to tax hikes. The remainder of the 2011 fiscal year — around five and a half months — thereore offers science agencies a short window of stability in which to take stock and prepare for renewed turmoil.

BASIC RESEARCH

Although few agencies would call themselves "winners" in the new budget, the sense of a narrow escape must have been palpable at the Department of Energy's Office of Science. Best known for funding high-energy and nuclear physics, the office had been targeted for an 18% cut relative to 2010 by House Republicans (see 'The arduous journey of the 2011 US budget'), a measure that would have meant thousands of lay-offs at national labs. In the end, the office came out of the deal with a cut of just over 2.1%. "It's a reasonable budget," says Pier Oddone, director of Fermilab in Batavia, Illinois. Oddone says that if the cut is applied evenly to all of the Office of Science's funding, Fermilab should still be able to continue to run its particle accelerator, ▶

➲ NATURE.COM
For more on the US budget wars, visit: go.nature.com/eg2no0

the Tevatron, until the end of this year.

Last week the White House conceded in a statement that the promised 10-year budget doubling for the physical sciences would probably not materialize. The America COMPETES Act, re-authorized last December, had called for the boost, relative to 2007 baselines, for the Office of Science, the National Science Foundation (NSF) and the National Institute of Standards and Technology (NIST). Investment would still be “strong”, it said.

Anita Jones, a computer scientist at the University of Virginia in Charlottesville, is unconvinced. Jones, who served on the National Academy of Sciences panel that called for the budget doubling in the influential report *Rising Above the Gathering Storm*, says that with the current budget, “Americans will cede the discovery of new ‘innovation territories’ to other nations”.

On the surface, the NIST seems to have been hit hard, with a 12.5% cut relative to 2010 levels. But its Scientific and Technical Research and Services, the account that covers the agency’s labs, was reduced by just 1.4%. The remaining cuts will be made in extramural research, development and construction. “I’d say the feeling is a mixture of relief that the NIST labs escaped larger cuts hitting other parts of the government, and disappointment that the increases promised by the America COMPETES act have failed to materialize,” says one NIST scientist who is not authorized to speak on policy.

BIOMEDICAL RESEARCH

At the National Institutes of Health (NIH), a cut of just over 1% amounts to a loss of $323 million, including $50 million that Congress has specified will come out of funding for buildings and facilities on the agency’s campus in Bethesda, Maryland. The rest will be distributed proportionately across the NIH’s 27 institutes and centres, and the office of the director.

“The bottom line for the NIH is: it could have been worse,” says Jennifer Zeitzer, the director of legislative relations at the Federation of American Societies for Experimental Biology in Bethesda, Maryland.

THE ARDUOUS JOURNEY OF THE 2011 US BUDGET (FIGURES IN US$ MILLION)

Agency	2010 enacted	2011 request	2011 — after proposed cut*	2011 enacted
National Institutes of Health (NIH)	31,089	32,089	30,519	30,766
NASA	18,725	19,000	18,421	18,448
Environmental Protection Agency (EPA)	10,281	10,020	7,283	8,812*
National Science Foundation (NSF)	6,873	7,424	6,597	6,807
Centers for Disease Control and Prevention (CDC)	6,475	6,342	5,419	5,733
Department of Energy Office of Science (DOE OS)	4,964	5,121	4,010	4,858
National Oceanic and Atmospheric Administration (NOAA)	4,748	5,555	4,631	4,597
Food and Drug Administration (FDA)	2,364	2,508	2,123	2,447*
US Geological Survey (USGS)	1,112	1,133	1,105	1,084
National Institute of Standards and Technology (NIST)	857	922	700	750

*Estimated

A notable absence from the new bill was any funding for the Cures Acceleration Network, an NIH programme designed to speed new treatments to the clinic, authorized in the landmark 2010 health-reform law. NIH director Francis Collins is hoping that the network will underpin the National Center for Advancing Translational Sciences that he intends to launch in October. He and the administration are asking for $100 million for the network in next year’s budget. However, the absence of any specified fund for it in 2011 does not bode well for funding levels in 2012.

“In that year and beyond, the threat is much greater,” says David Moore, director of government relations at the Association of American Medical Colleges in Washington DC.

The short-term reality is harsher for the Centers for Disease Control and Prevention (CDC), in Atlanta, Georgia. The portion of its budget controlled by Congress will be cut this year by 11.5%, or $742 million. (Some spending by the agency, for instance on vaccine programmes, is automatic and not under Congressional control.)

Donald Hoppert, director of government relations at the American Public Health Association in Washington DC, says that his group is “deeply disappointed” with the cuts to the CDC. “[They] will jeopardize our efforts to address the growing burden of chronic disease and injury, infectious disease and environmental hazards in communities across the nation. Any short-term savings are far outweighed by the long-term costs to the public’s health,” he says.

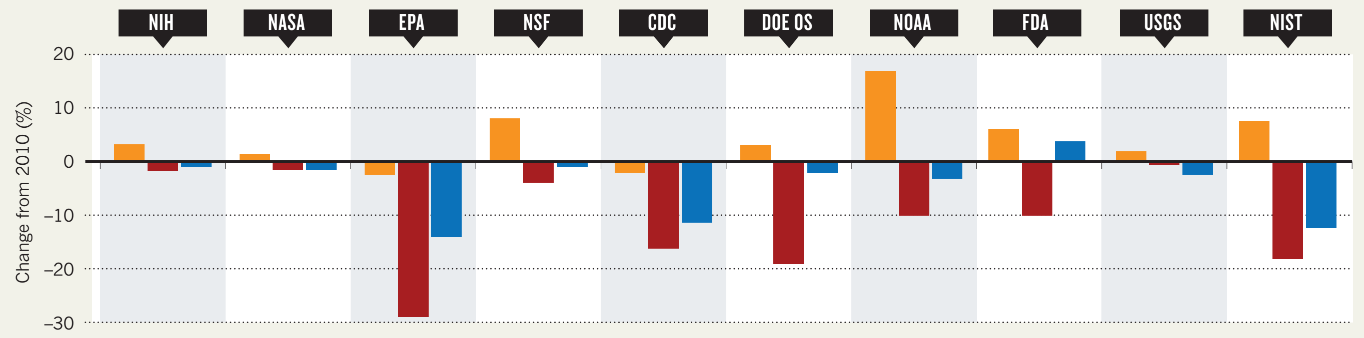
REGULATORY SCIENCE

During the budget debate, officials at the Food and Drug Administration (FDA) wondered whether the agency would receive enough money this year to begin to fulfil the requirements of the Food Safety Modernization Act. The act, which was signed into law on 4 January, gives the agency significant new powers to inspect both domestic and imported food products. House Republicans had proposed cutting the FDA’s budget by $241 million, which would have been a major setback for the act. Instead, the final deal gives the FDA an increase of about $83 million, more than half of which is allocated to food safety.

In the realm of environmental regulation, Senate Democrats and the White House successfully fended off Republican efforts to hobble the Environmental Protection Agency by restoring nearly half of the $3 billion in proposed cuts, and keeping

SPLITTING THE DIFFERENCE

The 2011 budget request made by President Barack Obama increased funding for most of the science agencies compared with 2010 levels. The final deal represents a compromise between those requests and the drastic cuts passed by the Republican-led House of Representatives.



cuts for science-and-technology research programmes at 4% relative to 2010. A Republican demand to strip the agency of its authority to regulate greenhouse-gas emissions was also removed from the final agreement. "On climate, it is a pretty strong victory for the administration," says David Goldston, director of government affairs at the Natural Resources Defense Council in Washington DC. Although the administration gave some ground, Goldston says, the bill no longer represents an outright assault on environmental policies. "The broad attack is not there any more," he says. The bulk of the cuts to the EPA — nearly $1 billion — have been made in grants to states to develop clean water-related infrastructure. The bill also cuts $112.5 million from grants to help state regulatory agencies prepare for new federal air-quality and climate regulations.

ENERGY, EARTH AND SPACE

The bill passed last week will reduce government funding for research on energy efficiency and renewable energy by $438 million, but that's modest compared with the much larger cut proposed by House Republicans. Similarly, the energy department's Advanced Research Project Agency-Energy will receive $180 million — far short of Obama's original $300-million request, but more than triple what the Republicans allocated in February.

At the National Oceanic and Atmospheric Administration (NOAA), the bill prevents money from being spent on the agency's Climate Service, which the administration has proposed as a central clearing house for climate data and longer-term forecasting. It also flatlines funding for weather- and climate-monitoring activities under the Joint Polar Satellite System. On 13 April, NOAA director Jane Lubchenco warned a Senate subcommittee that delays caused by the funding shortfall could jeopardize forecasting ability.

At NASA, a 1.5% cut will allow the space agency to proceed with plans to develop its human-spaceflight infrastructure, including a multipurpose crew vehicle and a heavy-lift rocket, to which the new spending bill allocates a minimum of $1.2 billion and $1.8 billion, respectively. NASA's science programme, which includes the over-budget James Webb Space Telescope (the launch of which has now been delayed until 2018), receives a robust $5 billion in the spending bill. But for NASA, as for other science agencies, the new budget leaves longer-term issues unsolved. With the final shuttle flight scheduled for June, the space agency faces a major transition this year — to a hiatus in which the nation is unable to fly its own astronauts. Future budgets will determine the length of that gap. ■

Additional reporting by Adam Mann and Gwyneth Dickey Zakaib

FUNDING

Canadian research shift makes waves

Agency's focus on industry-driven projects raises concerns that basic science will suffer.

BY HANNAH HOAG

Canada's largest research entity has a new focus — and some disaffected scientists. On 1 April, the National Research Council (NRC), made up of more than 20 institutes and programmes with a total annual budget larger than Can$1 billion (US$1 billion), switched to a funding strategy that downplays basic research in favour of programmes designed to attract industry partners and generate revenue. Some researchers suggest that the shift is politically driven, because it brings the agency into philosophical alignment with the governing Conservative Party of Canada, which is in the middle of an election campaign.

The change was announced in a memo from NRC president John McDougall on 2 March, and involves the transfer of authority over 20% of the agency's research funds and the entire Can$60-million budget for large equipment and building costs to the NRC's senior executive committee, which will direct it towards research with a focus on economic development, rather than pure science. Until now, individual institutes have had authority over research spending. McDougall wrote that in future, 80% of the research budget will be centralized, with "curiosity and exploratory activities" to be funded by the remaining 20%.

In Canada, most funding for academic researchers flows through agencies other than the NRC. However, with 4,700 scientists, guest researchers, technologists and support staff pursuing specialities from astrophysics to plant biotechnology at its institutes, the NRC plays a vital part in the nation's scientific community, as a generator of original research and a service provider to government and industry. The shift away from basic science "weakens" the NRC's labs, because they "are required to bridge two cultures — the basic and applied", says John Polanyi, a Nobel laureate and a chemist at the University of Toronto.

But in a follow-up memo on 24 March, McDougall said "most 'researcher directed' and basic work is now carried out in academic institutions. Duplicating the efforts of universities at NRC doesn't make much sense."

Four proposed 'flagship programmes' described in the original memo, each with a marketable outcome, provide a glimpse at the direction the agency has in mind. They include developing a strain of wheat resilient to environmental stress; improving the manufacture of printable electronics; increasing domestic production of biocomposite materials; and using algae to soak up carbon dioxide emissions from industry. NRC researchers have expressed concern that jobs and programmes that do not fit with the new agenda are at risk. The agency declined to comment.

Tom Brzustowski, who studies commercialization of innovation at the University of Ottawa, says that the adjustment to the NRC's focus will support areas that have been weak. "By focusing on the flagship programmes there is still room to do the whole spectrum of research. It's a good strategic move," he says.


NRC CANADA

"Duplicating the efforts of universities at NRC doesn't make much sense."
John McDougall

But the news has rekindled anxiety over how Canada's government has been directing science funding — criticisms that have grown sharper as the federal election on 2 May approaches.

On 22 March, the government presented its 2011 budget, which offered modest increases to the federal research councils, but did not make up for cuts in 2009 (see *Nature* **457**, 646; 2009). The budget also included multimillion-dollar investments in neuroscience and physics. Few question the quality of work that such investments would produce, but critics say that the government is exerting too much control over the country's research, rather than allowing peer review to guide funding.

"It's risky to divert funds away from the granting councils, but the government does it because it looks politically better for them," says Robert Dunn, associate director of scientific affairs at the Montreal Neurological Institute. "Peer review is the very best mechanism to ensure that the limited research resources we have are allocated to the best researchers and projects." ■


THE YALE DAILY NEWS

Fellow students hold a vigil for Yale undergraduate Michele Dufault.

HIGHER EDUCATION

A death in the lab

Fatality adds further momentum to calls for a shake-up in academic safety culture.

BY RICHARD VAN NOORDEN

In the early hours of 13 April, undergraduate students working at Yale University's Sterling Chemistry Laboratory made a shocking discovery. There in the lab's machine shop was the dead body of 22-year-old undergraduate student Michele Dufault, her hair tangled in a lathe. She had apparently died of asphyxiation in an accident described by Richard Levin, president of Yale in New Haven, Connecticut, as a "true tragedy".

Within days, federal health and safety officials had started to investigate. Details are scarce, but it is already clear that Dufault was not inexperienced with the equipment; she had taken a training course and had used the lathe safely many times before, according to fellow physics student Joe O'Rourke. She was, however, working late at night and probably alone (a speculation that Yale would not confirm) — circumstances that were not unusual at the machine shop, says O'Rourke.

Around the United States, laboratory directors and safety officers immediately checked their own policies on working practices in machine shops. But the accident has also heightened wider concerns about the ever-present tension between research freedom and safe working conditions in academia. And it underscores the slow pace of change since another high-profile laboratory fatality led to similar soul-searching less than three years ago.

In late 2008, 23-year-old research assistant Sheharbano Sangji sustained horrific burns in a lab fire at the University of California, Los Angeles (UCLA), and died of her injuries 18 days later. Sangji's death — in very different circumstances from Dufault's — resulted in federal fines for the university and a rapid toughening of safety policies there. On 30 March, UCLA unveiled its latest safety initiative: a new Center for Laboratory Safety, which is billed as the first in the United States to measure the effectiveness of safety policies and develop ways to improve scientists' approach to safety. More widely, Sangji's accident acted as a lightning rod for demands to improve standards across the United States.

Yet for all this attention, health and safety experts say that they have not seen a significant shift in the behaviour of bench scientists or the attitudes of lab heads, who are in the best position to improve safety culture. "It's very difficult to change principal investigators' attitudes," says James Gibson, UCLA's director of environmental health and safety. All too often, researchers in laboratories around the country still work alone, and without proper supervision or protection. "In many cases, academic freedom is more important than safety," says Jim Kaufman, president of the Laboratory Safety Institute in Natick, Massachusetts.

CHEMICAL FOCUS

Although such concerns apply across academia, accidents in chemistry laboratories have drawn the most scrutiny in recent years. A year after Sangji's death, Preston Brown, a graduate student in chemistry at Texas Tech University in Lubbock, lost three fingers of his left hand during a dangerous experiment. Brown was grinding up chunks of nickel hydrazine perchlorate — using a hundred times the recommended amount — when it detonated.

Unusually, the US Chemical Safety Board (CSB) — a body that usually investigates large industrial accidents such as refinery explosions — stepped in. For the first time ever, it said it would review academic laboratory safety. At an August 2010 meeting of the American Chemical Society in Boston, Massachusetts, CSB chairman Rafael Moure-Eraso said that the board had gathered media reports of around 120 university chemistry laboratory accidents since 2001, and concluded that "safety practices at US universities leave a lot to be desired".

Chemistry labs have been a particular focus of concern because the most dangerous procedures in other sciences tend to have more detailed safety protocols, says Peter Reinhardt, head of environmental health and safety at Yale University. "Using radioactive materials or biological materials is much more stringently regulated," he told *Nature* (speaking before Dufault's accident). "The big gap is hazardous chemicals in laboratories."

Rick Danheiser, an organic chemist at the Massachusetts Institute of Technology in Cambridge who chairs the chemistry department's health and safety committee, agrees that some labs' safety standards are too lax. But "there are chemistry departments with very strong safety programmes, and there's a whole range of laboratories in between", he says.

Neal Langerman, who runs the consulting company Advanced Chemical Safety, based in San Diego, California, is more strident about the extent of the problem. "I have come to the conclusion that most academic laboratories are unsafe venues for work or study," he wrote in a 2009 opinion column in the *Journal of Chemical Health and Safety*. He now says that, despite the recent accidents, he has not noticed a significant change in chemists' attitude to safety.

US scientists are undeniably much safer today than before swathes of occupational health legislation arrived in the 1970s, along with a new watchdog agency, the federal Occupational Safety and Health Administration (OSHA). Further improvements came in

1991, when OSHA stipulated that each chemistry lab should prepare a 'chemical hygiene plan' — effectively a handbook detailing safety protocols and emergency procedures — although these requirements are rarely enforced by inspections.

The US Bureau of Labor Statistics says that the rate of recordable incidents in scientific research and development services has fallen from 2.1 per 100 full-time employees in 2003 to 1.2 in 2009. But the government does not track major accidents or near misses specifically in academic laboratories. "Anecdotally, most people agree that university labs have more frequent and more frequently serious accidents than industry," says Dorothy Zolandz, director of the National Academies Board on Chemical Sciences and Technology.

Researchers often point out that industry is in a better position than academia to keep safety standards high because it has a clear hierarchy of power, fewer inexperienced students, and accountability to management. One of the clearest difference lies in lone working: surveys by the American Chemical Society last year suggest that 70.5% of faculty and 52.1% of graduate students often or occasionally work alone in laboratories, something that is forbidden in industry.

Safety officers and experienced chemists say that good laboratory safety relies on far more than regular inspections. What's key, says Tom Welton of Imperial College London, is that the group's research leader accepts unequivocal responsibility for the safety of everyone doing science in the laboratory, building a culture where researchers instinctively have safety foremost in their minds.

Evidence presented at a US National Research Council meeting on laboratory safety in November 2010 backs up Welton's point. Ron Zanoni, manager of occupational safety at international chemicals giant Arkema, based in Colombes, France, showed a 2004 survey that found case injury rates ranging from 7.8 to 0.8 per year at Arkema's various US sites. The differences correlated well with working relationships and top-down leadership engagement at different sites, Zanoni says. Improving safety leadership at labs with poor records had reduced injury rates by 2007.

As UCLA has found, it can be hard to change researchers' mindsets, even after a death on campus. Over the past two years, the university has ramped up laboratory safety regulations, training and inspections. But Nancy Wayne, a physiology professor on the board of the new laboratory safety centre, says researchers at UCLA do not always appreciate the tougher regime, sometimes seeing environmental inspectors as 'police', rather than partners in improving standards. "Changing the culture is really going to be a long-term challenge," says Gibson. Some professors, he says, have even questioned the need for flame-resistant lab coats — a bitter irony given the circumstances of Sangji's accident.

She was using a syringe to draw reactive *t*-butyl lithium from a bottle when it burst into flames, setting her clothes alight. She was not wearing a lab coat. Since then, the California Division of Occupational Safety and Health has agreed fines with UCLA of around US$70,000 for safety violations. Sangji's supervisor, Patrick Harran, declined *Nature*'s request for an interview. The Los Angeles district attorney is still reviewing Sangji's case, and has not yet decided whether to press criminal charges against either Harran or UCLA. If this resulted in a conviction, "the rules change completely right then and there", says Langerman. "All of a sudden, if you hurt somebody badly, you may face felony charges."

"Changing the culture is really going to be a long-term challenge."

In the United Kingdom, the threat of legal action has proved to be a powerful incentive for change. Around 25 years ago, an explosion in a chemistry laboratory at Sussex University in Brighton shot a piece of metal into a student's abdomen. The student eventually recovered, but the government's Health and Safety Executive prosecuted Sussex University for negligence. The episode had a profound effect on safety standards in Britain, says Welton. Today, British researchers are required to write down risk assessments before every experiment, something that is not required in the United States.

"I think that it will take a professor being punished, perhaps unfairly, to really engender change on the part of academia overall," says chemical-safety blogger Chemjobber, an industrial synthetic organic chemist in the United States. But funding agencies could also play a part. The CSB, for example, is considering recommending that grant applications should contain specific safety-training requirements.

"I think in the long run, the CSB recommendations, a possible new OSHA lab standard, and input from the American Chemical Society will result in a modification of the regulatory climate," says Langerman. But scientists should not wait for those changes before taking the initiative on safety, he adds.

"Members of the academic community have unique freedoms that are denied to industry," he says. "They then have a unique responsibility to behave in a manner that supports the freedom they are given." ■ SEE EDITORIAL P.259

FACEBOOK

Sheharbano Sangji (left) and Michele Dufault.

TUMOUR BIOLOGY

Cancer theory faces doubts

A leading explanation for how disease migrates falls short on clinical evidence.

BY HEIDI LEDFORD

It can't be easy to stand in front of hundreds of colleagues and tell them that ten years of research has led them in the wrong direction. But at the annual meeting of the American Association for Cancer Research (AACR) earlier this month, pathologist David Tarin did just that, by challenging a leading hypothesis on how cancers metastasize.

Tarin, from the University of California, San Diego, questioned the idea that cancer cells break free and migrate to new sites in the body by reverting to a state resembling the mobile cells in the developing embryo. Lab evidence for the hypothesis has raised hopes of some day blocking metastasis, which in 90% of cancer deaths is what ultimately kills the patient. But Tarin, along with some other cancer biologists, argues that no one has seen the process in action in human cancers.

Most cancers occur in the sheets of epithelial cells that line organs. Epithelial cells are normally immobile, but during embryonic development some start to produce proteins associated with motility and shut down production of proteins that glue the cells together. This transforms the cells into more mobile 'mesenchymal' cells that migrate to their correct locations in the embryo. If the same epithelial-to-mesenchymal transition (EMT) takes place in cancers, it could explain how tumour cells detach from their neighbours and enter the bloodstream to seed a new tumour (see 'Hypothesis under fire').

"It's a nice concept," says Pierre Savagner, a cancer researcher at the Montpellier Cancer Research Institute in France. "It makes things relatively simple to understand."

Although controversial at first, the EMT hypothesis was bolstered by experiments in mice showing that activation of EMT led to metastasis. OSI Pharmaceuticals, based in Melville, New York, is already pursuing EMT inhibitors as possible cancer treatments, as are several academic labs.

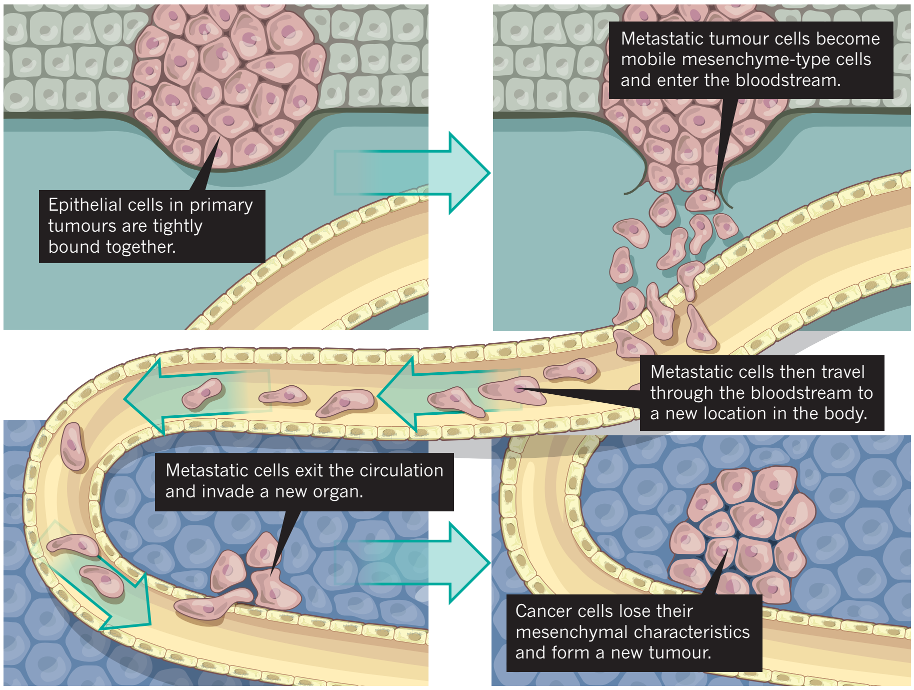
"EMT is really becoming a popular subject," says Shoukat Dedhar, a cancer researcher at the British Columbia Cancer Research Centre in Vancouver and a self-avowed convert to the hypothesis. "There are more and more data coming out for EMT's role in metastasis."

Yet sceptics remain. Tarin, among others, is concerned that the hypothesis has achieved popularity on the basis of results from cultured cells and animal models, without convincing evidence of its importance in metastasis in humans. He argues that pathologists have combed through millions of tissue sections from tumours without seeing cells in transition.

➲ NATURE.COM
For more on cancer and its prevention see:
go.nature.com/rtvavt

HYPOTHESIS UNDER FIRE

Reactivation of a key developmental pathway could explain how tumour cells invade distant tissues.



CONFLICTING IDEAS

Champions of the EMT–metastasis hypothesis, including leading cancer biologist Robert Weinberg at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, say that this may simply be because EMT is so transient — once a metastatic cell has invaded a new tissue, its mesenchymal features melt away. "When we examine a piece of tissue from a patient, we have just a snapshot of a moment," agrees cancer researcher Gianluigi Giannelli of the University of Bari Medical School in Italy. "We don't see the full movie."

Tarin dismisses this reasoning. "It's comparable to one telling you there are invisible aliens sitting in this room," he told the crowd at the AACR meeting, "but because you haven't used the right tools, you haven't seen them."

Others echo his concerns. Savagner points out that the proteins that might indicate EMT activation are also present during unrelated processes, such as programmed cell death. "People really want these cells to do an EMT and they push the concept too far," he says.

Even Dedhar worries about overreaching claims. For example, many papers report results from cultured cells, he says, without confirming that the cells are metastatic in live animals.

The real picture, Tarin says, may not involve a change in cellular identity. Instead, he and others suggest that metastasis occurs when mutations in cancer cells compromise cell–cell adhesion. Others suggest that cells break off from the tumour in clumps and travel in packs.

Weinberg says a key experiment that could resolve the controversy would be to track individual cancer cells from the time they break off from a tumour to the point at which they colonize a new organ. Such an experiment would be technically challenging in humans, he says. "One can study the primary tumour and the cells it dispatches into the circulation," Weinberg says. "But finding out what happens to these circulating tumour cells after they become lodged is very difficult."

"Whether this process is mandatory for metastasis to occur is still unclear," says Isaiah Fidler, a cancer biologist at the M. D. Anderson Cancer Center in Houston, Texas. "But EMT is not to be dismissed. In cancer, we can't dismiss anything." ■

WANO

Q&A Laurent Stricker

Nuclear safety chief calls for reform

Laurent Stricker, chairman of the World Association of Nuclear Operators (WANO), says that the disaster at Japan's Fukushima Daiichi nuclear plant should mark a turning point for an industry that many experts believe has become complacent about the safety of its reactors. Created in 1989, WANO is an international forum, headquartered in London, that brings together all nuclear power plant operators, along with governments and nuclear experts, to improve operational safety across the industry. Stricker is a nuclear engineer and former power plant director, and is also the senior adviser on nuclear affairs to the French utility company EDF.

Should Fukushima prompt WANO to change its remit?

Until now, WANO has addressed lessons learned from reactor operations, but not reactor design issues. I think in the future it should, in particular so that when operators modify their designs they draw more on analyses of past accidents.

It is not easy to designate one reactor design as safer than another. Rather, one must look at the particular case of each reactor's implementation, and its location. A reactor exposed to the threat of a tsunami doesn't face the same risks as a reactor of the same design elsewhere.

Population proximity is also very important. Japan, like many other countries, has several enormous nuclear sites near dense populations, so those demand even higher safety margins. After Fukushima, I believe that safety reviews should also consider the risk of accidents at several reactors at the same site at the same time. Often the current plans are only done for an accident in one reactor at a site.

We also need to be prepared for events exceeding what a reactor was designed to withstand, and to learn how best to cope with accidents such as a loss of electricity supply and cooling capacity, as happened at Fukushima Daiichi. That means having the right emergency procedures and equipment, and regular emergency drills, often involving the local population. Some countries do this very well; others do it much less, or not at all.

In October, WANO will bring together in China the chief executives of all the nuclear operators to discuss lessons learned from Fukushima, and any changes needed to WANO's mandate. WANO needs to be in a position to verify that every nuclear operating company has plans to cope with unforeseen accidents.

Has the industry been overconfident that a serious nuclear accident is now impossible?

Absolutely. I worry about overconfidence. People think we have good designs, we have good operators, we have good procedures and good safety authorities, so they think everything is fine.

Does the International Nuclear Events Scale distort the true safety record of the industry, with 'near misses' being registered as low-level incidents rather than potential disasters?

I think you are right. And it's true that the scale of severity is used in very different ways from one country to another. You also have differences in transparency from one country, and from one operator, to the next. At WANO, for example, we ask member companies to report incidents to us so that we can analyse them and share lessons from them. But between 5% and 7% of the power plants don't report any events in a given year. As an operator, I'm convinced that anyone running a nuclear power plant is bound to have something to report over the course of a year.

Could greater international oversight improve safety?

My point of view is that there are not enough plans in place to immediately help an operator in another country to cope with an accident.

Also, for countries that are relatively new to operating nuclear power plants, peer review before plant start-up is essential because serious accidents have often occurred in new reactors shortly after start-up. WANO sends teams of 20–25 engineers from other nuclear plants to review the functioning of the new plant for about three weeks and to provide a confidential report. The International Atomic Energy Agency has a similar programme that does five or six similar reviews per year; WANO has greater resources and conducts about 40 of these reviews a year. At our meeting in China, I will propose increasing their frequency.

I have also proposed that if operators fail to make progress on issues flagged by these reports as 'areas for improvement', then WANO should be authorized to dispense with its obligations of confidentiality.

If there is another major accident, is nuclear energy finished?

I fear so. As we have seen at Fukushima, an accident in one country has consequences for all nuclear operators elsewhere. ■

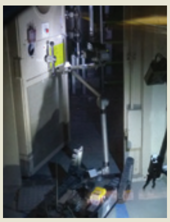
INTERVIEW BY DECLAN BUTLER

(Edited and translated from French.)

TEPCO

MORE ONLINE

TOP NEWS



Robots provide first glimpses into stricken Fukushima reactors go.nature.com/9jww7p

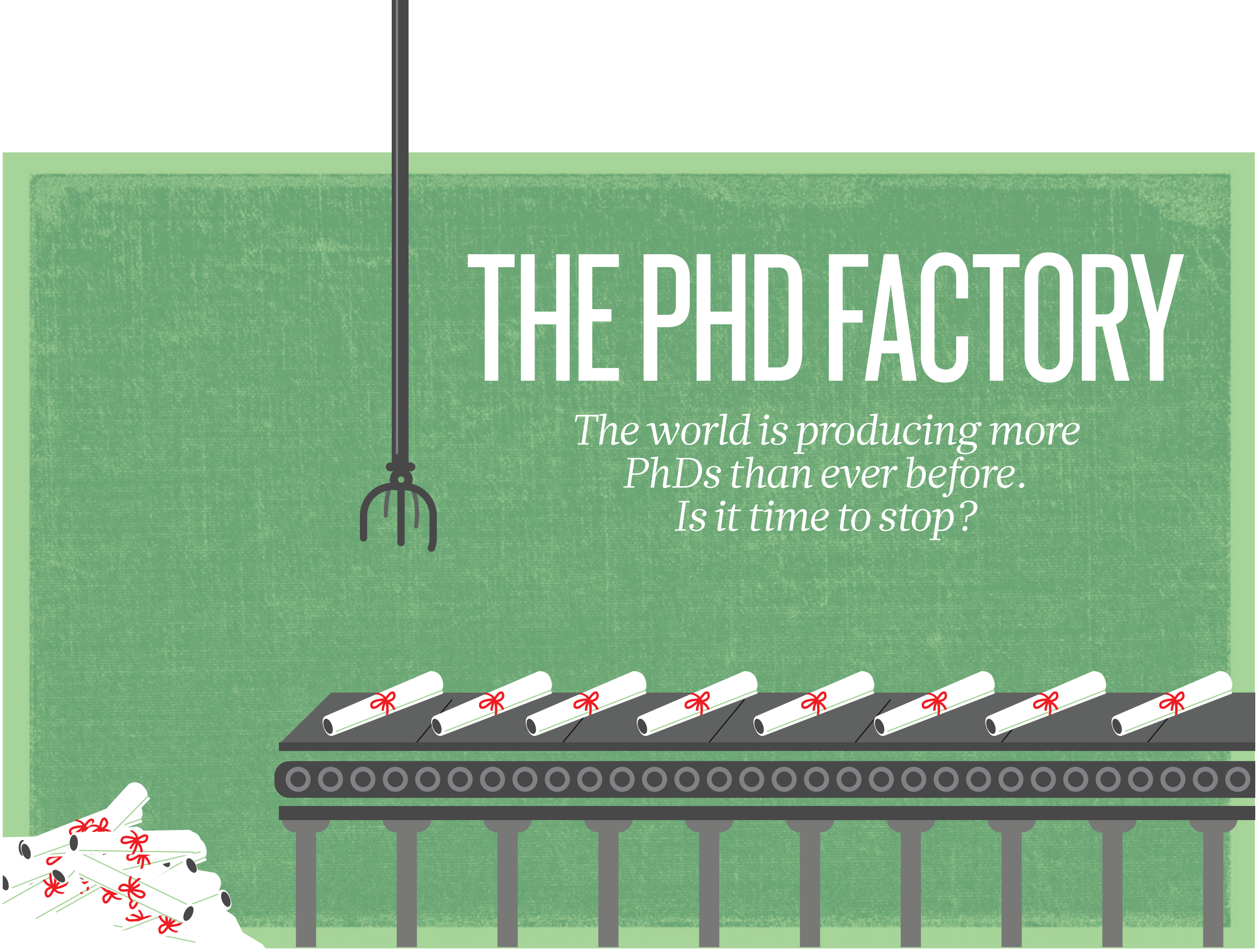
MORE NEWS

- Natural gas greenhouse-emissions study draws fire go.nature.com/akvfjj
- Is spate of large earthquakes statistical fluke or a sign of the times? go.nature.com/ocrgww
- How to choose which coral reefs to preserve go.nature.com/vtnqqr

CORRECTIONS

The news story 'China faces up to 'terrible' state of its ecosystems' (*Nature* **471,** 19; 2011) stated that more than 25% of China's grasslands have been lost in the past decade. The percentage should have been 2.5%

The Editorial 'Universal truths' (*Nature* **472,** 136; 2011) should have referred to Joseph Greenberg, not Joshua Greenberg.



THE PHD FACTORY

The world is producing more PhDs than ever before. Is it time to stop?

Scientists who attain a PhD are rightly proud — they have gained entry to an academic elite. But it is not as elite as it once was. The number of science doctorates earned each year grew by nearly 40% between 1998 and 2008, to some 34,000, in countries that are members of the Organisation for Economic Co-operation and Development (OECD). The growth shows no sign of slowing: most countries are building up their higher-education systems because they see educated workers as a key to economic growth (see 'The rise of doctorates'). But in much of the world, science PhD graduates may never get a chance to take full advantage of their qualifications.

In some countries, including the United States and Japan, people who have trained at great length and expense to be researchers confront a dwindling number of academic jobs, and an industrial sector unable to take up the slack. Supply has outstripped demand and, although few PhD holders end up unemployed, it is not clear that spending years securing this high-level qualification is worth it for a job as, for example, a high-school teacher. In other countries, such as China and India, the economies are developing fast enough to use all the PhDs they can crank out, and more — but the quality of the graduates is not consistent. Only a few nations, including Germany, are successfully tackling the problem by redefining the PhD as training for high-level positions in careers outside academia. Here, *Nature* examines graduate-education systems in various states of health.

JAPAN: A SYSTEM IN CRISIS

Of all the countries in which to graduate with a science PhD, Japan is arguably one of the worst. In the 1990s, the government set a policy to triple the number of postdocs to 10,000, and stepped up PhD recruitment to meet that goal. The policy was meant to bring Japan's science capacity up to match that of the West — but is now much criticized because, although it quickly succeeded, it gave little thought to where all those postdocs were going to end up.

> "EVERYONE TENDS TO LOOK AT THE FUTURE OF THE PHD LABOUR MARKET VERY PESSIMISTICALLY."

Academia doesn't want them: the number of 18-year-olds entering higher education has been dropping, so universities don't need the staff. Neither does Japanese industry, which has traditionally preferred young, fresh bachelor's graduates who can be trained on the job. The science and education ministry couldn't even sell them off when, in 2009, it started offering companies around ¥4 million (US$47,000) each to take on some of the country's 18,000 unemployed postdoctoral students (one of several initiatives that have been introduced to improve the situation). "It's just hard to find a match" between postdoc and company, says Koichi Kitazawa, the head of the Japan Science and Technology Agency.

This means there are few jobs for the current crop of PhDs. Of the 1,350 people awarded doctorates in natural sciences in 2010, just over half (746) had full-time posts lined up by the time they graduated. But only 162 were in the academic sciences or technological services,; of the rest, 250 took industry positions, 256 went into education and 38 got government jobs.

With such dismal prospects, the number entering PhD programmes has dropped off (see 'Patterns of PhD production'). "Everyone tends to look at the future of the PhD labour market very pessimistically," says Kobayashi Shinichi, a specialist in science and technology workforce issues at the Research Center for University Studies at Tsukuba University.

CHINA: QUANTITY OUTWEIGHS QUALITY?

The number of PhD holders in China is going through the roof, with some 50,000 people graduating with doctorates across all disciplines in 2009 — and by some counts it now surpasses all other countries. The main problem is the low quality of many graduates.

Yongdi Zhou, a cognitive neuroscientist at the East China Normal University in Shanghai,

identifies four contributing factors. The length of PhD training, at three years, is too short, many PhD supervisors are not well qualified, the system lacks quality control and there is no clear mechanism for weeding out poor students.

Even so, most Chinese PhD holders can find a job at home: China's booming economy and capacity building has absorbed them into the workforce. "Relatively speaking, it is a lot easier to find a position in academia in China compared with the United States," says Yigong Shi, a structural biologist at Tsinghua University in Beijing, and the same is true in industry. But PhD graduates can run into problems if they want to enter internationally competitive academia. To get a coveted post at a top university or research institution requires training, such as a postdoctoral position, in another country. Many researchers do not return to China, draining away the cream of the country's crop.

The quality issue should be helped by China's efforts to recruit more scholars from abroad. Shi says that more institutions are now starting to introduce thesis committees and rotations, which will make students less dependent on a single supervisor in a hierarchical system. "Major initiatives are being implemented in various graduate programmes throughout China," he says. "China is constantly going through transformations."

SINGAPORE: GROWTH IN ALL DIRECTIONS

The picture is much rosier in Singapore. Here, the past few years have seen major investment and expansion in the university system and in science and technology infrastructure, including the foundation of two new publicly funded universities. This has attracted students from at home and abroad. Enrolment of Singaporean nationals in PhD programmes has grown by 60% over the past five years, to 789 in all disciplines — and the country has actively recruited foreign graduate students from China, India, Iran, Turkey, eastern Europe and farther afield.

Because the university system in Singapore has been underdeveloped until now, most PhD holders go to work outside academia, but continued expansion of the universities could create more opportunities. "Not all end up earning a living from what they have been trained in," says Peter Ng, who studies biodiversity at the National University of Singapore. "Some have very different jobs — from teachers to bankers. But they all get a good job." A PhD can be lucrative, says Ng, with a graduate earning at least S$4,000 (US$3,174) a month, compared with the S$3,000 a month earned by a student with a good undergraduate degree.

"I see a PhD not just as the mastery of a discipline, but also training of the mind," says Ng. "If they later practise what they have mastered — excellent — otherwise, they can take their skill sets into a new domain and add value to it."

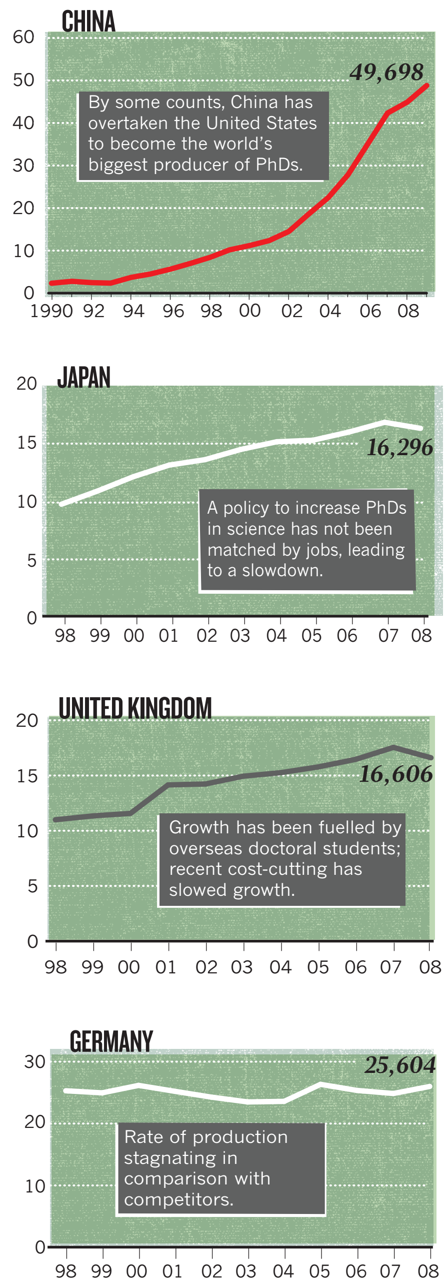
UNITED STATES: SUPPLY VERSUS DEMAND

To Paula Stephan, an economist at Georgia State University in Atlanta who studies PhD trends, it is "scandalous" that US politicians continue to speak of a PhD shortage. The United States is second only to China in awarding science doctorates — it produced an estimated 19,733 in the life sciences and physical sciences in 2009 — and production is going up. But Stephan says that no one should applaud this trend, "unless Congress wants to put money into creating jobs for these people rather than just creating supply".

The proportion of people with science PhDs who get tenured academic positions in the sciences has been dropping steadily and industry has not fully absorbed the slack. The problem is most acute in the life sciences, in which the pace of PhD growth is biggest, yet pharmaceutical and biotechnology industries have been drastically downsizing in recent years. In 1973, 55% of US doctorates in the biological sciences secured tenure-track positions within six years of completing their PhDs, and only 2% were in a postdoc or other untenured academic position. By 2006, only 15% were in tenured positions six years after graduating, with 18% untenured (see 'What shall we do about all the PhDs?'). Figures suggest

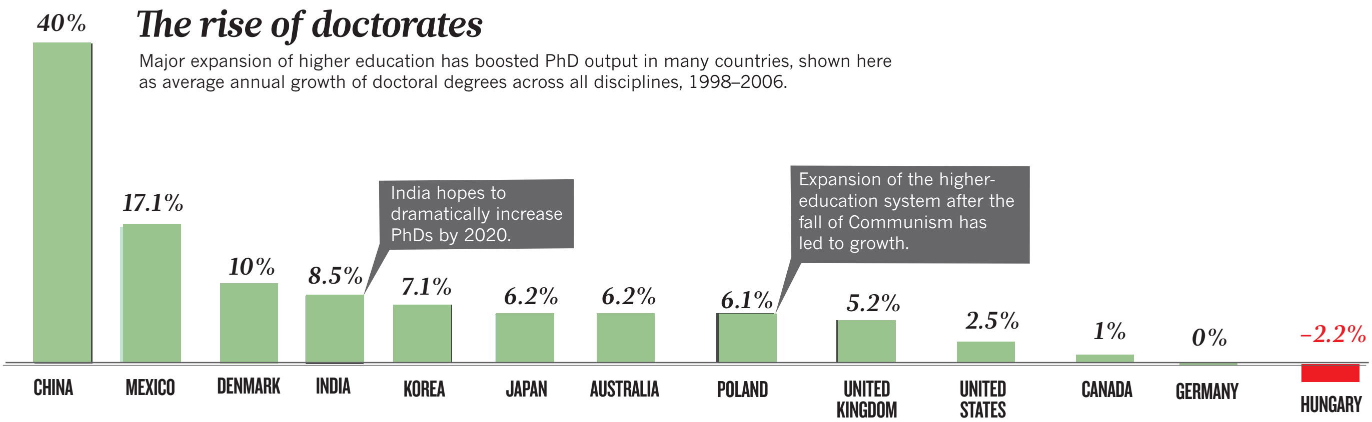
Patterns of PhD production

Trends in annual PhD graduation across all disciplines. All figures given in thousands of PhDs.



The rise of doctorates

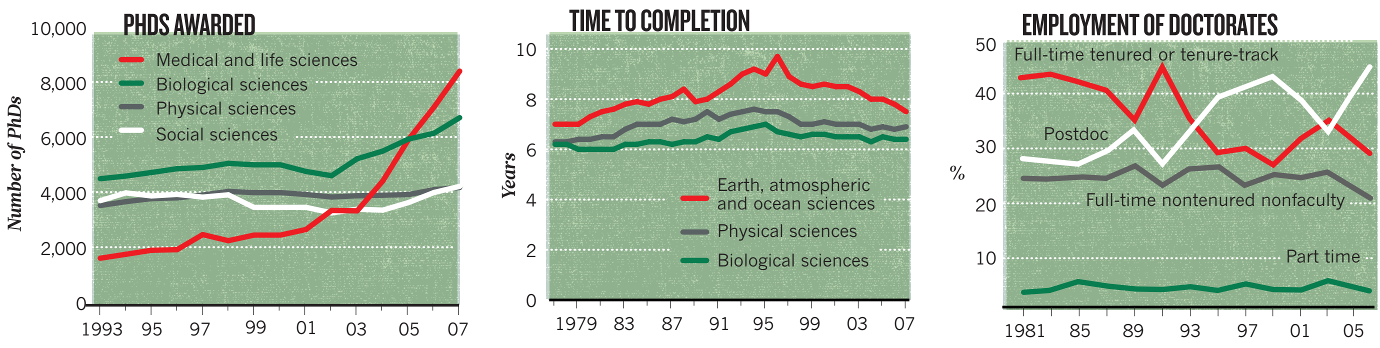
Major expansion of higher education has boosted PhD output in many countries, shown here as average annual growth of doctoral degrees across all disciplines, 1998–2006.



SOURCE: OECD/CHINESE MINISTRY OF EDUCATION

United States: What shall we do about all the PhDs?

The annual number of science and engineering doctorates graduating from US universities rose to almost 41,000 in 2007 (left), with the biggest growth in medical and life sciences. It took a median of 7.2 years to complete a science or engineering PhD (middle) — yet the proportion finding full time academic jobs within 1–3 years of graduating is dwindling (right).



SOURCE: SCIENCE AND ECONOMIC INDICATORS 2010

that more doctorates are taking jobs that do not require a PhD. "It's a waste of resources," says Stephan. "We're spending a lot of money training these students and then they go out and get jobs that they're not well matched for."

The poor job market has discouraged some potential students from embarking on science PhDs, says Hal Salzman, a professor of public policy at Rutgers University in New Brunswick, New Jersey. Nevertheless, production of US doctorates continues apace, fuelled by an influx of foreign students. Academic research was still the top career choice in a 2010 survey of 30,000 science and engineering PhD students and postdocs, says Henry Sauermann, who studies strategic management at the Georgia Institute of Technology in Atlanta. Many PhD courses train students specifically for that goal. Half of all science and engineering PhD recipients graduating in 2007 had spent over seven years working on their degrees, and more than one-third of candidates never finish at all.

Some universities are now experimenting with PhD programmes that better prepare graduate students for careers outside academia (see page 280). Anne Carpenter, a cellular biologist at the Broad Institute of the Massachusetts Institute of Technology (MIT) and Harvard University in Cambridge, Massachusetts, is trying to create jobs for existing PhD holders, while discouraging new ones. When she set up her lab four years ago, Carpenter hired experienced staff scientists on permanent contracts instead of the usual mix of temporary postdocs and graduate students. "The whole pyramid scheme of science made little sense to me," says Carpenter. "I couldn't in good conscience churn out a hundred graduate students and postdocs in my career."

But Carpenter has struggled to justify the cost of her staff to grant-review panels. "How do I compete with laboratories that hire postdocs for $40,000 instead of a scientist for $80,000?" she asks. Although she remains committed to her ideals, she says that she will be more open to hiring postdocs in the future.

➲ NATURE.COM
Tell us what you think about the future of PhDs:
nature.com/phdfuture

GERMANY: THE PROGRESSIVE PHD

Germany is Europe's biggest producer of doctoral graduates, turning out some 7,000 science PhDs in 2005. After a major redesign of its doctoral education programmes over the past 20 years, the country is also well on its way to solving the oversupply problem.

Traditionally, supervisors recruited PhD students informally and trained them to follow in their academic footsteps, with little oversight from the university or research institution. But as in the rest of Europe, the number of academic positions available to graduates in Germany has remained stable or fallen. So these days, a PhD in Germany is often marketed as advanced training not only for academia — a career path pursued by the best of the best — but also for the wider workforce.

> "THE RELATIVELY LOW INCOME OF GERMAN ACADEMIC STAFF MAKES LEAVING THE UNIVERSITY AFTER THE PHD A GOOD OPTION."

Universities now play a more formal role in student recruitment and development, and many students follow structured courses outside the lab, including classes in presenting, report writing and other transferable skills. Just under 6% of PhD graduates in science eventually go into full-time academic positions, and most will find research jobs in industry, says Thorsten Wilhelmy, who studies doctoral education for the German Council of Science and Humanities in Cologne. "The long way to professorship in Germany and the relatively low income of German academic staff makes leaving the university after the PhD a good option," he says.

Thomas Jørgensen, who heads a programme to support and develop doctoral education for the European University Association, based in Brussels, is concerned that German institutions could push reforms too far, leaving students spending so long in classes that they lack time to do research for their thesis and develop critical-thinking skills. The number of German doctorates has stagnated over the past two decades, and Jørgensen worries about this at a time when PhD production is growing in China, India and other increasingly powerful economies.

POLAND: EXPANSION AT A COST

Growth in PhD numbers among Europe's old guard might be waning, but some of the former Eastern bloc countries, such as Poland, have seen dramatic increases. In 1990–91, Polish institutions enrolled 2,695 PhD students. This figure rose to more than 32,000 in 2008–09 as the Polish government, trying to expand the higher-education system after the fall of Communism, introduced policies to reward institutions for enrolling doctoral candidates.

Despite the growth, there are problems. A dearth of funding for doctoral studies causes high drop-out rates, says Andrzej Kraśniewski, a researcher at Warsaw University of Technology and secretary-general of the Polish Rectors Conference, an association representing Polish universities. In engineering, more than half of students will not complete their PhDs, he says. The country's economic growth has not kept pace with that of its PhD numbers, so people with doctorates can end up taking jobs below their level of expertise. And Poland needs to collect data showing that PhDs from its institutions across the country are of consistent quality, and are comparable with the rest of Europe, says Kraśniewski.

Still, in Poland as in most countries, unemployment for PhD holders is below 3%. "Employment prospects for holders of doctorates remain better than for other higher-education graduates," says Laudeline Auriol, author of an OECD report on doctorate holders between 1990 and 2006, who is now analysing doctoral-student data up to 2010.

Still, a survey of scientists by *Nature* last year showed that PhD holders were not always more satisfied with their jobs than those without the degree, nor were they earning substantially more (see 'What's a PhD worth?').

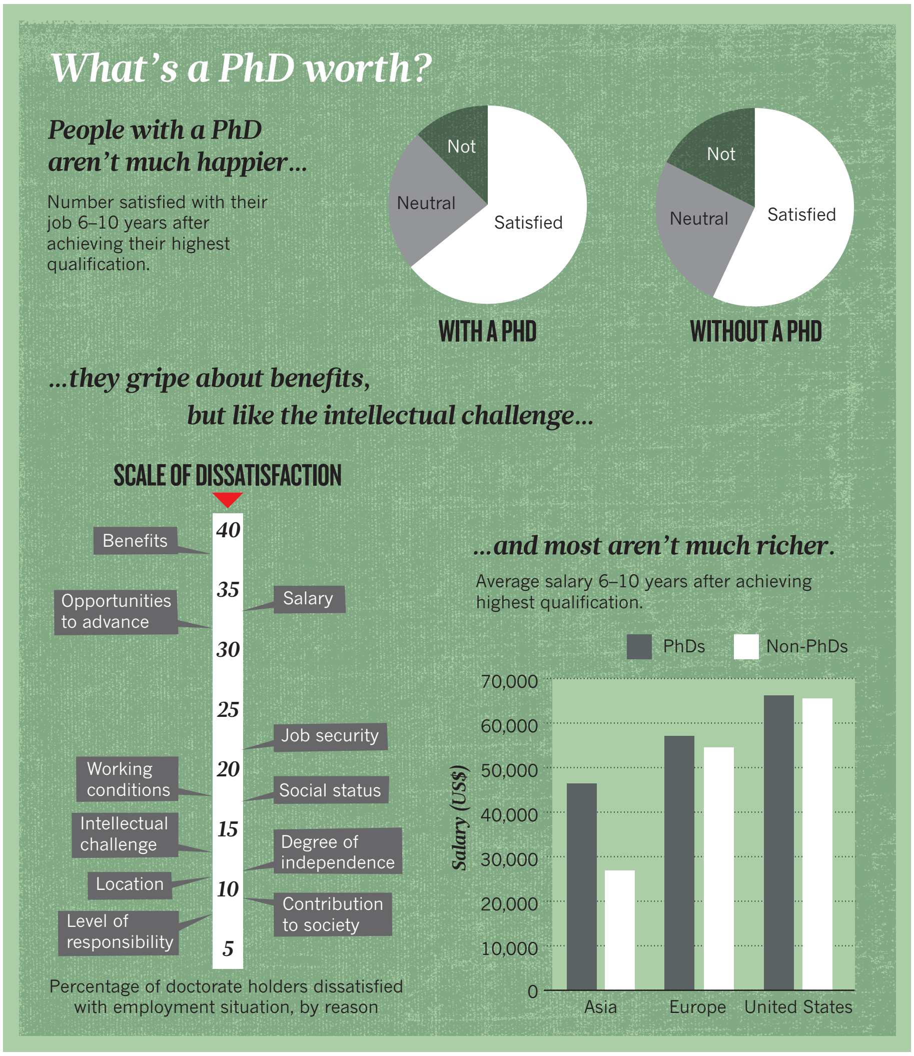
EGYPT: STRUGGLE TO SURVIVE

Egypt is the Middle East's powerhouse for doctoral studies. In 2009, the country had about 35,000 students enrolled in doctoral programmes, up from 17,663 in 1998. But funding has not kept up with demand. The majority comes through university budgets, which are already strained by the large enrolment of students in undergraduate programmes and postgraduate studies other than PhDs. Universities have started turning to international funding and collaborations with the private sector, but this source of funding remains very limited.

The deficit translates into shortages in equipment and materials, a lack of qualified teaching staff and poor compensation for researchers. It also means that more of the funding burden is falling on the students. The squeeze takes a toll on the quality of research, and creates tension between students and supervisors. "The PhD student here in Egypt faces numerous problems," says Mounir Hana, a food scientist and PhD supervisor at Minia University, who says that he tries to help solve them. "Unfortunately, many supervisors do not bother, and end up adding one more hurdle in the student's way."

Graduates face a tough slog. As elsewhere, there are many more PhD holders in Egypt than the universities can employ as researchers and academics. The doctorate is frequently a means of climbing the civil-service hierarchy, but those in the private sector often complain that graduates are untrained in the practical skills they need, such as proposal writing and project management. Egyptian PhD holders also struggle to secure international research positions. Hana calls the overall quality of their research papers "mediocre" and says that pursuing a PhD is "worthless" except for those already working in a university. But the political upheaval in the region this year could bring about change: many academics who had left Egypt are returning, hoping to help rebuild and overhaul education and research.

Few PhDs are trained elsewhere in the Middle East — less than 50 a year in Lebanon, for example. But several world-class universities established in the oil-rich Gulf States in recent years have increased demand for PhD holders. So far, most of the researchers have been 'imported' after receiving their degrees from Western universities, but Saudi Arabia and Qatar in particular have been building up their infrastructure to start offering more PhD programmes themselves. The effect will be felt throughout the region, says Fatma Hammad, an endocrinologist and PhD supervisor at Al-Azhar University in Cairo. "Many graduates are now turning to doctoral studies because there is a large demand in the Gulf States. For them, it is a way to land jobs there and increase their income," she says.



SOURCE: OECD, NATURE SALARY & CAREER SURVEY 2010

INDIA: PHDS WANTED

In 2004, India produced around 5,900 science, technology and engineering PhDs, a figure that has now grown to some 8,900 a year. This is still a fraction of the number from China and the United States, and the country wants many more, to match the explosive growth of its economy and population. The government is making major investments in research and higher education — including a one-third increase in the higher-education budget in 2011–12 — and is trying to attract investment from foreign universities. The hope is that up to 20,000 PhDs will graduate each year by 2020, says Thirumalachari Ramasami, the Indian government's head of science and technology.

Those targets ought to be easy to reach: India's population is young, and undergraduate education is booming (see *Nature* **472,** 24–26; 2011). But there is little incentive to continue into a lengthy PhD programme, and only around 1% of undergraduates currently do so. Most are intent on securing jobs in industry, which require only an undergraduate degree and are much more lucrative than the public-sector academic and research jobs that need postgraduate education. Students "don't think of PhDs now, not even master's — a bachelor's is good enough to get a job", says Amit Patra, an engineer at the Indian Institute of Technology in Kharagpur.

Even after a PhD, there are few academic opportunities in India, and better-paid industry jobs are the major draw. "There is a shortage of PhDs and we have to compete with industry for that resource — the universities have very little chance of winning that game," says Patra. For many young people intent on postgraduate education, the goal is frequently to go to the United States or Europe. That was the course chosen by Manu Prakash, who went to MIT for his PhD and now runs his own experimental biophysics lab at Stanford University in California. "When I went through the system in India, the platform for doing long-term research I didn't feel was well-supported," he says. ■

Reporting by **David Cyranoski, Natasha Gilbert, Heidi Ledford, Anjali Nayar** *and* **Mohammed Yahia.**

ILLUSTRATION BY: OLIVER MUNDAY

RETHINKING PHDS

Fix it, overhaul it or skip it completely — institutions and individuals are taking innovative approaches to postgraduate science training.

BY ALISON MCCOOK

"Most of them are not going to make it." That was the thought that ran through Animesh Ray's mind 15 years ago, as he watched excellent PhD students — including some at his own institution, the University of Rochester in New York — struggle to find faculty positions in academia, the only jobs they had ever been trained for. Some were destined for perpetual postdoctoral fellowships; others would leave science altogether.

Within a few years, the associate professor was in a position to do something about it. A stint in a start-up company in California had convinced him that many PhD graduates were poor at working in teams and managing shifting goals, the type of skills that industrial employers demand. So he started to develop a programme that would give students at Keck Graduate Institute (KGI) in Claremont, California, these skills. "I was determined not to have to keep watching scientists struggle to find the jobs they were trained to do."

Ray is one of a number of researchers and administrators who are attempting to reshape graduate training. They want to save young scientists from falling into the postdoc holding pattern or taking jobs below their station. Here, *Nature* presents five approaches to shaking up the hallowed foundations of academia. They range from throwing scientists deep into independent study, to going interdisciplinary, to forgoing the PhD altogether.

1 JUMP IN AT THE DEEP END

For Michael Lenardo, a molecular immunologist at the US National Institutes of Health (NIH) in Bethesda, Maryland, the thought process went like this: When too many scientists are looking for too few academic positions, PhD programmes need to admit the students most likely to succeed, and provide them with all the skills they'll need. And neither the United States nor the United Kingdom seemed to be getting the mix exactly right.

In the United Kingdom, PhD students are given independence early, and degrees rarely last more than 4 years. But not all institutions require that students publish a first-author paper, which Lenardo sees as a drawback. US science degrees often do require first-author papers, but have ballooned to more than 7 years in duration.

In 2001, Lenardo created a new degree programme, called the NIH Oxford-Cambridge Scholars Program, that would combine the best elements of each system for a cadre of truly elite students. It admits just 12 of the 250–300 applicants per year. Independence is stressed — students devise and write their own project plan, begin their thesis work immediately, and skip the uniform coursework — but they must meet requirements such as authoring papers.

Students split their time between the United States and the United Kingdom, and have at least two mentors, one in each country (and often in different disciplines). Because no adviser has full control, students learn how to operate independently, says Lenardo. Travelling to another country reinforces that autonomy, and ensures that the students work with the best people in their field, he says.

In the ten years since the programme's inception, more than 60 students have graduated, taking slightly more than 4 years apiece. They published an average of 2.4 first-author papers out of their PhD research. Eighty percent of graduates are still in academia, and half a dozen are already working as principal investigators.

Ambika Bumb, now a postdoc at the National Cancer Institute in Bethesda, spent her PhD developing a nanoparticle with magnetic, optical and nuclear properties that might one day aid in imaging tumours and delivering targeted therapies. She finished in just three years, had four advisers in two countries and received training in engineering, immunology, radiochemistry and radiology. She published at least four scientific papers and one review article from her PhD research, and she is now applying for faculty positions.

Developing independence is a crucial step to becoming an investigator, says Richard Hetherington, a postgraduate-skills development coordinator at Newcastle University, UK. "Having that will make them stronger when they get to the end," he says. But a lack of structure and core coursework could leave some students unprepared, says Nathan Vanderford, who manages a grant and manuscript development office at the University of Kentucky in Lexington, and has written about career issues in science. "I don't see that you'd get the depth of the history [of science], and the central core principles, strictly in a lab setting." Some students may struggle.

2 FORGET ACADEMIA

Ray's experiences encouraged him to think more about non-academic training for PhDs. Many institutes, including KGI, had already embraced Professional Science Master's (PSM) programmes as a way to stock the ranks of industry and keep training scientists, but Ray found that these degrees could limit students' opportunities.

He watched as graduates of KGI's Master's of Bioscience often started as an assistant to a consultant, or a mid-level manager, then advanced from there. They did well, but typically remained in the management side of a company, separate from the science. So Ray worked with David Galas, a KGI co-founder, and Sheldon Schuster, the institute's president, to extend the PSM's reach and develop a PhD programme that would provide students with both industry know-how and technical research training.

To complete a PhD in Applied Life Sciences at KGI, students must first complete the master's course there, then spend three to four more years doing original research, with at least one adviser from industry. Eric Tan, the first graduate of the programme, spent his PhD at KGI developing a DNA chip that might have applications in diagnostics or assessing biological threats. He learned not only the scientific method, but also how to write a business plan and present it to venture capitalists, how to carry out market research and the ins and outs of patent legislation.

> "I WAS DETERMINED NOT TO HAVE TO KEEP WATCHING SCIENTISTS STRUGGLE."

Courses in marketing and communication are useful for any scientist, even those who stay in academia, says Vanderford. "Regardless of the career path a PhD would take, having those courses would be helpful."

Time will tell if it is working. Ray is inspired by the success of KGI's PSM programme, which has seen nearly all of its 300 graduates find jobs since it started in 2000. Since the PhD programme began in 2006, three students have earned their degrees, and each has found a job earning more than the median starting salary for the PSM students (US$73,000). It is a result that Ray calls "astounding".

Ray says he hopes that the rounded training will give his students the ability to manage scientists and interact with business people. "They can see and appreciate the big picture; at the same time, they are well-versed in the technological depth for which they will be valued."

But well-rounded students may have some dull edges, and Ray acknowledges that KGI cannot provide coursework in specific areas such as physical chemistry or cell biology. It will be an "ongoing process to try to figure out the balance between how much detailed science courses you need versus how much professional development you need", says Vanderford.

3 TRAMPLE THE BOUNDARIES

Marc Jacofsky was working on a PhD in physical anthropology at Arizona State University (ASU) in Tempe when his brother, an orthopaedic surgeon, told him about all the questions he wanted to investigate in movement and artificial joints. Jacofsky remembers interrupting his brother with a few suggestions: "He looked at me and he said, 'I thought you studied monkeys.'"

Jacofsky did study monkeys — but also engineering, mathematics, computer science, kinesiology and neurophysiology. He was enrolled in a new programme developed by ASU faculty members from a wide range of departments, an attempt to go beyond interdisciplinary studies and instead create entirely new disciplines.

Nearly every new PhD programme at ASU is designed to be "transdisciplinary", says Maria Allison, dean of the graduate college. Other examples include Human and Social Dimensions of Science and Technology, Biological Design and Urban Ecology. Some degrees involve more than 80 faculty members, because of the range of topics covered.

The initial funding for Jacofsky's programme, called Neural and Musculoskeletal Adaptations in Form and Function, and some of the other ASU degrees came from a National Science Foundation project known as IGERT, or Integrative Graduate Education and Research Traineeship. IGERT provides US$3-million 5-year grants to US institutions to develop programmes that help students to gain career skills and tackle real-world problems.

Since 1998, the IGERT programme has funded nearly 5,000 graduate students. An independent survey found that IGERT students are better able than their non-IGERT peers to work in multidisciplinary teams and to communicate with non-experts, without sacrificing expertise in their chosen area. There is even some indication that IGERT graduates have an easier time finding a job.

Similar interdisciplinary programmes are starting up elsewhere. The Canadian government has an initiative called the Collaborative Research and Training Experience Program, and a new PhD course in Bangalore, India, trains engineers, chemists, computer scientists and physicists in interdisciplinary life sciences, teaching them to use the tools of physical science to tackle biological problems. Started around five years ago by physicists at the National Centre for Biological Sciences, the Interdisciplinary Biology, or iBIO, programme has graduated eight students. Two are already tenure-track faculty members.

It is good to expose trainees to different fields, but specialization is still important, says Hetherington. The purpose of a PhD is to provide a "deep understanding of a specific area". Even cross-disciplinary research consists of scientists who contribute specific skills from their particular fields, he says.

Broadening the scope of a programme has advantages, however. It teaches students about their options. Jacofsky had entered his degree thinking he would one day teach

university-level anthropology. Instead, he is vice-president of research and development at the at the Center for Orthopedic Research and Education, or CORE Institute, in Phoenix, Arizona, co-owned by his brother. Jacofsky studies biomechanics and gait before and after orthopaedic procedures. "If I'd done a traditional anthropology degree, I think there's an incredibly small chance I'd be working in industry."

4 GET IT ONLINE

Some potential postgraduate students do not have the flexibility to commit to full-time studies, or to travel to a lab. Online training aims to fill this gap and provide more individuals with appropriate training, even at the PhD level.

Rana Khan started teaching an online course initially out of curiosity — she didn't understand how it would work. "I was fascinated by the whole idea," she says. "How do you do it?"

At the time, she was a postdoc at the US Department of Agriculture, investigating how to make soya beans more resistant to pathogens. She wanted teaching experience, and saw a job listing at the University of Maryland University College in Adelphi.

The job was to teach part of an online biotechnology Master's degree. The college had set up an online classroom, where Khan posts weekly lectures, and students are required to complete assignments and participate in discussions throughout the week. At least once a day, Khan checks in, answering students' questions. At the end of the programme, students do an online internship, in which they do group projects for real companies — investigating, for example, potential competitors with a new technology — and submit 100–200 page reports. There is no lab component, but there could be, says Khan, who directs the programme, now a PSM: students could simply work at a nearby lab and submit their data online, she says.

The college's programme has been around since 2001 and now graduates approximately 50 students a year. Roughly 10% live outside the United States. That's a big advantage of online degrees, Khan notes — some of her current students are members of the military, stationed in Afghanistan and Iraq.

One graduate is Kyle Retterer, who started a PhD in physics. After realizing he didn't want to spend years focusing on a narrow area in semiconductors, he abandoned academia. When he began to miss research, he looked for programmes that tackled cutting-edge problems and let him do what he had always loved — analyse huge amounts of data.

His mother had completed two online degrees in information technology and is now a vice-president at Nasdaq, so he saw the potential in distance learning. He graduated in two years, and two months later had a job at GeneDx, a clinical genetic-testing company in Gaithersburg, Maryland, analysing data from multi-gene tests. He now makes three to four times what he was making as a graduate student. "I feel like I'm in pretty good shape."

Even a PhD is possible from a distance. The Open University, which is headquartered in Milton Keynes, UK, now has about 40 part-time science PhD students. They work locally, conducting research at a local astronomy lab, for example, then are expected to check in every two weeks via Skype — or sometimes in person — with supervisors, usually at the university's main campus. "That can be just as rewarding" as having a supervisor on-site, says James Bruce, who manages the university's science PhD students.

Online PhDs are a rarity, but that could change, speculates Hetherington. Science isn't done in isolation, he says, so degrees in which students work alone and simply check in with a mentor won't teach them about managing relationships with mentors and peers. However, future tools could make it easier for students to interact with others remotely, better preparing them for being collaborative researchers, he says. "It will become increasingly more possible to do it."

➲ NATURE.COM
Join the discussion on the future of the PhD:
go.nature.com/nezhym

5 SKIP THE PHD

Some are choosing to forgo the PhD altogether. Deanna Pickett had always expected to get a PhD, maybe in engineering or environmental chemistry. That changed last year, during her final year as an undergraduate in chemistry at the College of Wooster in Ohio. Paul Edmiston, a chemistry professor, asked her to help him investigate the properties of a new material that absorbed contamination from drinking water. It was real work that had an immediate impact; she loved it.

So when she later visited a potential graduate school, she was unimpressed. The prospect of years of more theoretical work, when she was already doing field research, was unappealing. When Stephen Spoonamore, the chief executive and co-founder (along with Edmiston) of the company ABSMaterials in Wooster, asked her to continue her work after she graduated, she changed her plans. "It is just a little more fulfilling next step of my life than going to do another five years of research on another topic."

Pickett's opportunity is unusual, perhaps more so now than ever before. Academia and industry have such a rich choice of PhD graduates for jobs that those without PhDs need not apply. "There is currently an ample supply of highly skilled people on the market," says David Harwell, assistant director of career management and development at the American Chemical Society in Washington DC. In some fields, such as bioinformatics, simple on-the-job training can sometimes suffice, but even then scientists generally need a PhD to advance. "Anyone can cite examples of non-PhD bioinformaticists who have made really major contributions, but few of these people have taken on the full range of responsibilities typically reserved for PhD investigators," says Maynard Olson, a genomics researcher at the University of Washington in Seattle.

ONLINE PHDS ARE A RARITY, BUT THAT COULD CHANGE.

ABSMaterials is one of the few exceptions — mostly because Spoonamore believes that PhDs "have got the wrong training". Spoonamore says that he often pays undergraduates "about the same" as PhDs, and promotes them just as easily. He himself founded 13 technology companies without finishing an undergraduate degree, the first at the age of 18 with funding from his lawn-mowing business. "I will always have a preference for an incredibly smart, top-of-their-class undergraduate student in chemistry. Every time."

In her second day on the job, Pickett gave a presentation to a group of entrepreneurs, and a week later, had to develop a pilot plan to clean up a site in Ohio that had been contaminated with trichloroethylene. She says she probably does many things a PhD graduate would do. "I do feel like I've skipped a step," she says.

But she knows she might not get as many responsibilities if she decided to change companies. For this reason her colleague, Laura Underwood, has decided to pursue a PhD after working with ABSMaterials for 3 years. Underwood, who has a similar background to Pickett, was the company's first employee, with huge responsibilities — running a manufacturing facility, overseeing conference planning and managing a lab. Without a PhD, she fears it might be hard to find the same kinds of opportunities elsewhere. But she's glad she worked for a while before going to grad school. "If you go straight into a PhD, something that sounds great in a lab may be kind of underwhelming when you get into the field."

Alison McCook *is a freelance writer in Philadelphia, Pennsylvania.*

COMMENT

DESIGN Evolution of conservation practice traced in charity logos **p.287**

CONSERVATION A history of the WWF at 50 makes sobering reading **p.290**

FILM Isabella Rossellini on animal behaviour **p.294**

OBITUARY Jurg Tschopp, immunologist who changed patients' lives **p.296**

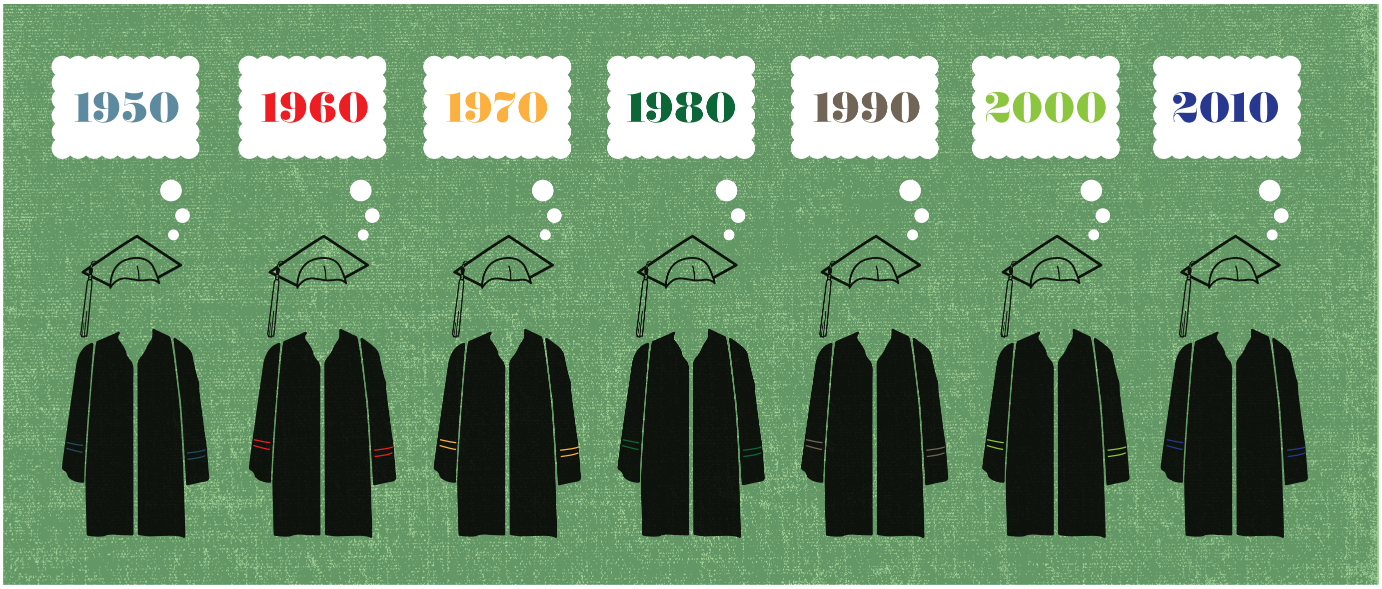


ILLUSTRATION BY OLIVER MUNDAY

Seven ages of the PhD

Scientists share memories of doing doctorates in different decades, disciplines and locations, from the hunt for the structure of DNA to deciphering the human genome.

RAYMOND GOSLING
1950s: The age of formality

Emeritus professor, University of London

PhD in structural analysis of DNA (King's College London, 1954).
Looking back on my PhD at King's College London, I realize that I was fortunate to experience the feudal process of research in a large department dominated by an enthusiastic professor, John Randall. Although I, like every other student, had to register for a PhD with Randall, I was free to collaborate with other senior scientists in the lab.

I worked closely with Maurice Wilkins and later Rosalind Franklin, under whose direction I pursued X-ray diffraction studies of the sodium salt of DNA. One of the highlights of my PhD was the now iconic photo 51, a diffraction photograph I took, which clearly showed the helical structure of the molecule (see image). I will always remember the moment I first saw that beautiful double diamond pattern.

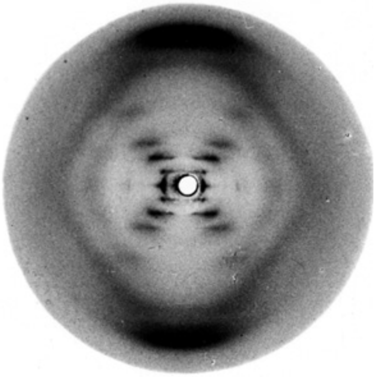
Of my two mentors, Maurice was friendly and determined but very shy. Rosalind had a more combative style and taught me the value of being a devil's advocate, so challenging and defending our developing ideas. However, I was unable to ease the tensions that developed between Maurice and Rosalind.

Randall's biophysics unit was a wonderfully energetic place to work. However, in those days relationships between staff and students were rather formal. All the men wore ties with their white lab coats, and the senior common room at King's was for men only. However, all the women scientists, including Rosalind, were an integral part of the lab's efforts.

Randall was exceptional because he found financial backing for all the members of his lab. The only other comparable example that I have experienced is the work at Cold Spring Harbour Laboratory in New York directed by James Watson. These days, students are not always so fortunate and have to spend time and energy seeking funding.

My advice to would-be doctoral students is to seek out a dynamic, flourishing research group. Do not start until you have agreed with your supervisor on a beginning, middle and an end to your project. A good PhD often raises more questions than it solves, so you should not be surprised if your work changes direction.

The University of London's advice to examiners states that PhD students should identify a distinct and new contribution to science, not just a review of the literature. Since those rather heady days of the 1950s, I have sometimes wondered whether this always happens. Some will say I was lucky ▶



R. GOSLING/KING'S COLLEGE LONDON ARCHIVES

Gosling's PHD photo revealing structure of DNA.

▸ to be in the right place at the right time, and I would agree, but it is also important to have high standards, something I learned from Randall.

The less formal relationships between students and professors today probably improve communications. I have since directed the research of many PhD students, and have come to know them all personally, something that was just not possible with Randall. He told me that he didn't want to see my PhD until it was submitted — I can't imagine that happening nowadays.

CHERYLL TICKLE
1960s: The age of independence

Emeritus professor, University of Bath, UK

PhD in cell biology (University of Glasgow, 1970).

As an undergraduate, I became fascinated by the arrangement of cells and how it is key to understanding embryology and tissue biology. So I applied to do a PhD with Adam Curtis at the University of Glasgow in Scotland. He had recognized that cell adhesion was central to 'sorting out', the process by which a mix of different cells stratifies into regions of cells of the same type.

The highlight of my PhD was managing, at the end, to publish a paper in the *Journal of Embryology and Experimental Morphology*, with a statistician. The biggest pressure I felt during my PhD was the responsibility for the data I produced. I realized for the first time that others might base their own work on what I had found and that there was nowhere I could go and look up the 'correct' answers. Part of my anxiety stemmed from the fact that although I had done a lot of practical work, I had not carried out a research project during my undergraduate degree.

The person who influenced me most during my doctorate was my supervisor. Curtis took me on not knowing much about me, suggested the project and then left it to me to work out the details. This meant that my progress was slow and sometimes painful, but also that I learned for myself how to set about things and master new techniques. Fortunately other students and staff in the department were encouraging and supportive. One of my most important practical lessons I learned the hard way: you can never record too much detail about your experiments. I was perplexed because I was suddenly getting different results. I later realized that I had switched to using a different shaker to bring cells together. This changed the results completely and I hadn't initially been recording which shaker I had used.

Nowadays, PhDs are much more structured. Students are not given as free a rein as I was, nor are they allowed to make as many mistakes. There is a greater emphasis on acquiring data. Students also often work with others rather than alone. These differences reflect the changes over the past 40 years in the way in which research is carried out, and its growing pressures.

STEVE W. RUNNING
1970s: The age of innocence

University of Montana, Missoula

PhD in forest ecology (Colorado State University, 1979).

I finished my PhD in 1979, just before personal computers arrived. So it was written on a typewriter, with 53 hand-drawn figures. The subject of this labour of love was inducing water stress on 13 pine trees by cutting their leaves off and measuring desiccation responses. Looking back now, my PhD research seems highly esoteric. The work built fundamental understanding of leaf-scale physiology but had no policy relevance. The only practical value was in understanding why your Christmas tree turns brown. I think how innocent we all were then, doing weird science and looking for cheap beer.

After my PhD I got a tenure-track professorship at the University of Montana in Missoula, teaching tree biology. I never imagined being interviewed by national journalists, or getting hate mail from the public. As a lead author for the Intergovernmental Panel on Climate Change's Fourth Assessment Report, I have now experienced both. What happened? The policy relevance of my work is to be found at larger scales than my 13 trees. For example, it can apply to the hydrology management of a river basin. So what began as curiosity-driven research into water stress on trees evolved into global analyses of terrestrial carbon sources and sinks. And questions in forest ecology are now motivated by policy and economics, including carbon credits, carbon offsetting and biofuels.

The modern PhD student needs to be much more policy aware, because society has many environmental problems to solve, and not much time. There is a huge need for better quantification of ecosystem services, and connecting sustainable ecosystem principles with slow-growth economics.

Public outreach has become essential. Young faculty members are often devoting time to blog discussions, work that needs to be valued better by academia. During my PhD, I had no training in public speaking, yet it may now be the most important thing I do. Above all, I think that this next generation of scientists will have no choice but to pursue research aimed at saving the planet from catastrophe. I remind my public audiences that Earth doesn't need us — the cockroaches will gladly take over if we flame out.

S. RUNNING

A far cry from hate mail: Steve Running induces water stress in pine trees in 1972.

TAO IMAGES/ALAMY

Yao Tandong spent a month during his PhD on Glacier No. 1 in the Tianshan mountains of northwest China.

YAO TANDONG
1980s: The age of internationalism

Head of the Institute of Tibetan Plateau Research, Beijing

PhD in glaciology (Institute of Geography, Beijing, 1986).

I graduated from Lanzhou University in China's Gansu Province in 1978. At that time, the postgraduate system in China was immature. During the Cultural Revolution (1966–76), the whole education system was broken, and it was only after 1978 that the degree system was restored. There were probably only a few hundred PhD students in China. I was one of the first to study for a master's degree in physical geography at Lanzhou, and because no institution in my province could grant a PhD I had to travel to Beijing to study at the Institute of Geography.

The prestigious institute was part of the Chinese Academy of Sciences. In the 1980s, science was held in the highest esteem in China, and you had to get the highest marks in very strict exams to get into a PhD programme. However, there were still few senior scientists who were qualified by the Chinese government to supervise PhD students. I was lucky in that my PhD adviser was the famous Chinese glaciologist Shi Yafeng.

It was also popular at that time to study abroad. I passed all my exams to teach English as a foreign language and planned to go to the United States, but my adviser asked me to stay in China. He said I could go abroad for short courses, but he wanted me to do most of my research in China. It turned out to be the right decision for me because I spent a whole month with him at the Glacier No. 1 research station in the Tianshan mountains in northwest China. This glacier is famous worldwide with scientists and tourists, and it is shrinking at an accelerating rate. It supplies waters to the city of Urumqi and is the world's closest glacier to a metropolis.

Shi Yafeng was already in his sixties, but very dedicated to his work — a dedication that inspired me for the rest of my career. During the month we spent on the glacier he worked for every possible hour on the water-resource problem that was the subject of my PhD. While I was on the glacier I also got to know the famous glaciologist Lonnie Thompson from Ohio State University in Columbus. He was just starting his ice-core work in China and that meeting shaped the rest of my career. After I graduated in 1986, I spent three years abroad working with ice-core scientists in Grenoble, France, and in the United States.

I returned to China in 1989 and initiated ice-core studies on the Tibetan plateau. I have now supervised more than 25 PhD students in the same field, although the country's great passion for science has been replaced by a national passion for business.

I was lucky enough to have a good adviser and excellent collaborators, but I had to find collaborators outside China because there were so few scientists to work with at home. I always encourage my students to spend some time abroad. I learned important new fieldwork and lab techniques in the two months I spent in Alaska during my PhD. Scientifically, we still have a gap in China. Happily, almost all the students from my institute have returned to China from overseas, so everyone benefits.

➲ NATURE.COM
Comment on the future of the PhD at:
nature.com/phdfuture

ANDRAS DINNYES
1990s: The age of revolution

Professor of biotechnology, Szent Istvan University, Gödöllő, Hungary

Candidate of sciences in veterinary science (1995)

The year 1989 was a remarkable one for me and for Hungary. I obtained my veterinary diploma in Budapest, and my country saw the official end to four decades of Communism. Free multiparty elections produced a new political system and in principle a new era in Hungarian science. I was eager to start a doctoral programme, because I had already done three years of scientific research during my veterinary studies and had become a recipient of the Pro Scientia Gold Medal of the Hungarian Academy of Sciences in the first year it was awarded.

In later years this award would have been a ticket for automatic admission to a doctoral programme, but in 1989 I still had to take an entry exam. I was admitted to the Academy of Sciences for a three-year fellowship programme towards a 'candidate of sciences' degree. This was the old Soviet-style degree — usually obtained by scientists in their forties, and an entry into the closed Hungarian scientific community. By the mid-1990s, it was replaced by a Western-style PhD, aimed at younger students and opening up the entire system. But along the way the academy's fellowship programme was new and brave, breaking with the old gerontocracy, and offering a 'fast-track' doctorate for a few young scientists.

To my surprise, I found my activities tightly controlled by my supervisor. I had to spend one year doing veterinary obstetrics work for the university, unrelated to my fellowship research in embryo cryopreservation. So during the second year, I applied for and won a Fulbright Fellowship to the Smithsonian Institution in Washington DC. There, I spent 16 months with William Rall, the 'godfather' of embryo vitrification technology (which uses high dose chemicals to aid embryo cryopreservation without the ice causing damage).

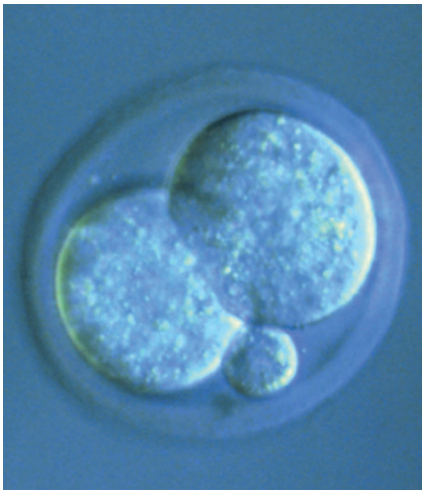
The fellowship changed my life. At first it was a culture shock to work for a strict American boss — David Wildt, the head of department — who demanded six days a week, ten hours a day. Rall provided me with training, advice and friendship and allowed me to work fairly independently. I gained self-confidence and became a workaholic. Soon I found international friends and enjoyed free museums and classical music concerts. The methods I learned determined the next ten years of my career as a scientist and inventor in embryo, ▶

gamete and stem-cell cryobiology.

In Washington, I produced sufficient data for my first publication and my thesis. Naively, I turned down an offer to do a PhD at Cornell University in Ithaca, New York, because I was so close to proudly completing my degree in Hungary.

I returned to Hungary in 1992 and submitted my completed thesis to the academy. My Hungarian supervisor thought it was premature, but I was stubborn and self-confident. Frustratingly, the scientific degree committee of the academy took more than two years to process my submission. So in the interim I did a one-year European Union postdoctoral fellowship in Belgium, for which the submitted thesis made me eligible. Finally, in 1995, at the age of 29, I was allowed to defend my thesis and became one of the youngest veterinarians to obtain the 'candidate of sciences' degree in Hungary.

After several positions abroad I am now back in a much-changed Hungary, working on genetic reprogramming, cloning and stem-cell research as a professor and chief executive of my own stem-cell company. I supervise six PhD students from Hungary and Thailand, who enjoy the reforms that came just a little too late for me. I always encourage my students to pursue international experience — which I can provide through the European Marie Curie projects I coordinate.



Andras Dinnyes worked on embryo preservation during his PhD in Hungary and the United States.

A. A. OSOWOLE
2000s: The age of perseverance

Department of Chemistry, University of Ibadan, Nigeria

PhD in inorganic chemistry (University of Ibadan, 2002).

I started my PhD on the kinetics of novel organometallic compounds at the University of Ibadan, Nigeria, in 1994, and completed it eight years later. After a year, a piece of kit that was crucial to my kinetics studies broke down and could not be repaired — the machine was ten years old and the manufacturer had stopped making it. Such technical difficulties are common in Nigeria because equipment is rarely well maintained.

I had to change my research topic to something I could do with the equipment available. So I began to investigate molecules containing both organic and inorganic components. This required the synthesis of novel metal complexes. The other researchers in the team encouraged me, gave me tips and taught me to use the equipment. A faculty member, Gabriel Kolawole, provided the expertise and a chemical to solve the hydrolysis problems I was facing.

Even with the right tools, the work was challenging. I did not isolate a single complex in the first year.

I also had three BSc students working with me, who needed to finish their dissertations within six months. The students were intended to facilitate my research, but they all needed close supervision. I spent about 35% of my time teaching and 65% on research. I was under enormous pressure: I worked virtually every day, about six hours on the bench as well as teaching and practical classes. By the end of 1998–99 year (the 1995–96 and 1996–97 sessions were cancelled due to industrial crises) I had managed to isolate about 100 metal complexes — a good result.

My next challenge was interpreting and writing up my results. In 1999, the University of Ibadan library had limited access to journals, the most recent ones dating to 1996. To get more up-to-date chemical papers I had to rely on colleagues abroad. Back then, things weren't as computerized. Furthermore, analysing my samples required equipment we did not have, necessitating sending my samples abroad for analyses. But I could not afford the fees.

I realized I would not be able to complete my PhD at Ibadan with these limitations. Again, help came from Kolawole, who suggested that I apply to the Third World Organization for Women in Science — now the Organization for Women in Science for the Developing World.

I was awarded a fellowship and used it in 2000 at the Indian Institute of Science in Bangalore. There I was able to complete the outstanding analysis, and learn porphyrin chemistry and various chromatographic techniques. I defended my thesis in 2002 and got about ten publications from it.

The lessons I learned from my doctorate which I try to pass on to my students are: be focused, persevere, work hard and be honest. Where there is a will, there is a way.

ERIKA CULE
2010s: The age of communication

Imperial College London

PhD in statistical genetics (Imperial College London, due for completion in 2012).

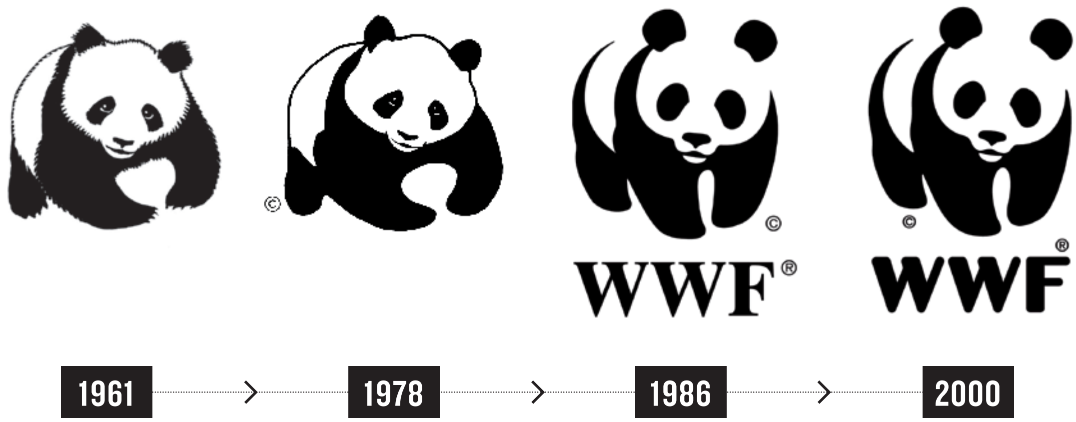
The first draft of the human genome was published when I was still at school. I was excited by the opportunities this milestone would offer, and back then I imagined that I would become a laboratory scientist. During my undergraduate degree, however, I found that I preferred analytical questions to cell and molecular biology. Today I am privileged to be working in a high-profile and rapidly evolving field — statistical genetics.

The biggest challenge of my PhD so far has been mastering the skills I need to do computational genetics, and building the confidence to use them effectively. The transition from undergraduate studies in biochemistry has been a steep learning curve. A master's in bioinformatics and theoretical systems biology enabled me to tackle programming and statistics and gave me some tools to get started.

Fortunately, I have had support. My late grandfather, himself an academic, helped me decide to study for a PhD and gave me practical advice. My elder sister is a recent PhD graduate, and she understands the ups and downs of the process. I also have an excellent relationship with my supervisor, who is actively interested in my work — I know that this is not always the case. A PhD can be an isolating experience, but it helps to work in a friendly and sociable department. Invaluable support also comes from my partner and friends who provide me with a perspective from outside academia.

Working in a computational discipline, and having an interest in writing, it seemed natural to start a blog. Blogging about my PhD (go.nature.com/zkcy5l) has helped me to connect with scientific bloggers from around the world who share their experiences and suggestions. I also hope that my blog writing will help me when it comes to writing the thesis.

As a new PhD graduate I will face a difficult climate in which research funding from the government is being squeezed and the academic job market is increasingly competitive. I hope that my PhD will equip me with specific and transferable skills — such as networking and presenting — that will give me options after I graduate, whether in academia or elsewhere. ■ SEE NEWS FEATURES P.276 & P.280



WWF

The art of conservation

As the World Wide Fund for Nature (WWF) turns 50, **Henry Nicholls** traces how the evolution of conservation practice has been echoed in the various incarnations of WWF's iconic pandas, and other conservation logos.

The World Wide Fund for Nature (WWF) came up with its famous panda logo 50 years ago. According to Max Nicholson, the mastermind behind the charity (then called the World Wildlife Fund) it was "one of the most valuable trademarks that has ever been devised, and it took about twenty minutes"[1].

It is natural that conservation organizations should borrow motifs from nature, but the motives that lead to the final designs of their logos are not always obvious. In 1961, as the WWF's founders mulled over the choice of their symbol in a plush town house in London's Belgravia, the most important consideration was that it should reproduce well on the organization's letterhead. With colour printing then out of the question for a fledgling charity, this narrowed the options to a shortlist of black-and-white species, and the popular panda emerged.

This mundane explanation is probably also behind several other famously two-toned conservation brands. One was the oryx chosen by Fauna and Flora International (FFI) in 1950 (Fig. 1k), another the avocet first used in print by the UK Royal Society for the Protection of Birds (RSPB) in 1966 (Fig. 1i). With advances in printing technology from the 1950s onwards, more colour — notably greens, frequently blues and sometimes browns — crept into conservation imagery. But monochromatic species continued to be popular. Birdlife International chose an arctic tern in 1991, its simple contrasting lines helping the organization to get noticed in our busy digital world (Fig. 1a).

➲ NATURE.COM
For the biodiversity special collection, visit:
go.nature.com/d1kgn9

Since 1961, just about every conservation brand has changed — including the WWF's panda. They have evolved in response to shifts in the media landscape, corporate life and conservation practice. This logo evolution, sometimes slow and incremental, sometimes rapid and radical, traces the story of how conservation charities have weathered the past half decade.

HOMESPUN LOOK

It is fair to say that 50 years ago most conservation outfits, including the WWF, were run by a close-knit core of passionate individuals operating on a shoe-string budget. It is no surprise, then, to find that conservation logos from this period were frequently homegrown. One option was to run a competition. This is how, in 1954, the International Union for the Conservation of Nature (IUCN) settled upon its 'flaming artichoke': an unspecified organic growth emerging from a cumbersome acronym (Fig. 1h). An alternative approach was to lean on an artistic friend. The Fauna Preservation Society (the forerunner of the FFI), for example, collared a lifetime member to ink out the black-and-white face of a gemsbok for the cover of the society's new journal. Better still was to call upon ornithological enthusiast, environmentalist and accomplished artist Peter Scott. He produced a sublime gannet for the British Trust for Ornithology (BTO) in the late 1940s (Fig. 1g), and a pair of forward-looking Bewick's swans for another British charity, the Wildfowl & Wetlands Trust (Fig. 1e).

What's striking about these early examples of conservation artwork is that many of the species on show were not in peril, probably because so little was known about population sizes and extinction risk in the mid-twentieth century. This was certainly the case for the giant panda. When the WWF alighted on its emblem in the early 1960s — also drawn by Scott — China's first dedicated census of the species was still more than a decade away. Even if the panda was having a hard time of it, as seems likely, this was not something the WWF chose to emphasize. Instead, they preferred to spin the idea that the panda "owes its survival to the sort of careful conservation which all wild creatures deserve"[2].

Soon, however, logos began to portray species that did have a clear conservation message. The dodo, the icon of extinction, was a perfect image for the Durrell Wildlife Conservation Trust, founded in 1963 to support Gerald Durrell's pioneering conservation-focused captive work at Jersey Zoo in the Channel Islands (Fig. 1d). There were also more upbeat emblems. The moving tale of Elsa the lioness (star of Joy Adamson's novel

FROM BACKROOM TO BOARDROOM

As conservation has changed, its logos have evolved, from literal, fine-art creations to abstract images that make only a passing reference to nature.

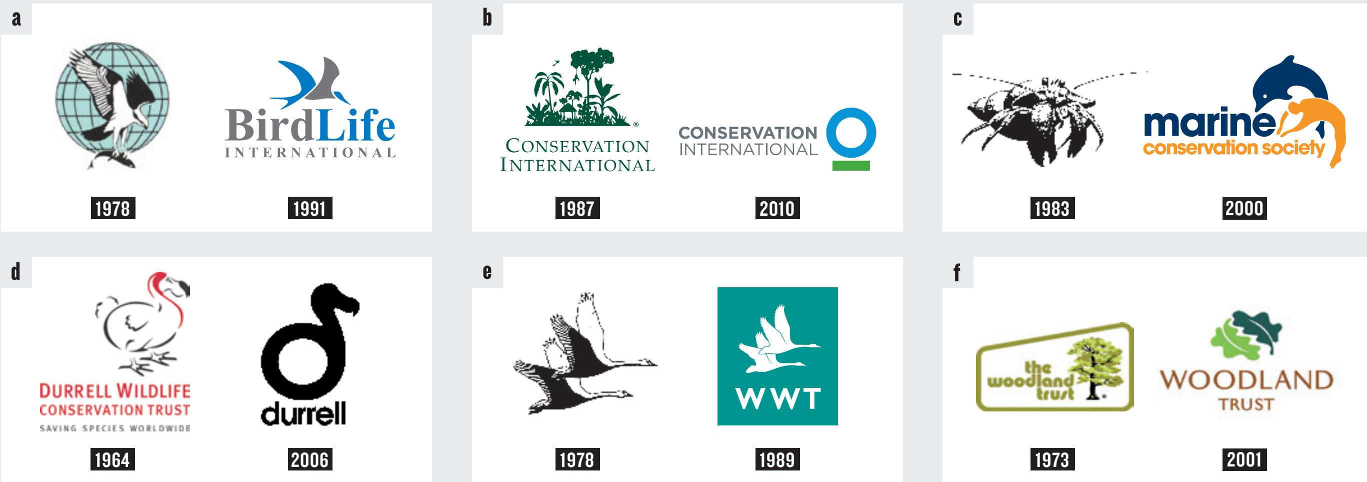


Figure 1 | Changing logos. a, Birdlife International. **b**, Conservation International. **c**, Marine Conservation Society. **d**, Durrell Wildlife Conservation Trust. **e**, Wildfowl & Wetlands Trust. **f**, The Woodland Trust. **g**, British Trust for Ornithology. **h**, International Union for the Conservation of Nature. **i**, Royal Society for the Protection of Birds. **j**, The Nature Conservancy. **k**, Fauna and Flora International. **l**, Friends of the Earth International.

Born Free (Collins and Harvill; 1960) and its 1966 Hollywood adaptation and memorable soundtrack) made her an exemplary face of the Born Free Foundation, established in 1984 to campaign against zoos and promote conservation in the wild. The RSPB's success in recreating the habitat suitable for breeding avocets in Britain during the 1940s made this species an obvious choice as an emblem (in addition to its being monochrome).

The FFI's common gemsbok (*Oryx gazella*) took on new meaning in the wake of Operation Oryx, a 1962 effort to save the critically endangered Arabian oryx (*Oryx leucoryx*) in which the FFI played a leading role[3]. For all but the most nerdy naturalists able to tell their *O. gazella* from their *O. leucoryx*, the FFI's logo suddenly became a celebration of a conservation triumph. Several artistic mutations later and the metamorphosis is nearly complete, with the FFI's latest logo (launched in 2010) looking more like the Arabian oryx than the gemsbok. It also has a spray of vegetation thrown in as an acknowledgement of the organization's commitment to flora.

CORPORATE CONSERVATION

By the 1980s, the conservation movement had gathered significant momentum. Between 1961 and 1981, for example, the WWF had raised some US$55 million in support of 2,800 conservation and education projects worldwide[4]. With awareness and funding on the increase and organizations growing from local concerns to national and international affairs with burgeoning overheads, it was inevitable that the bigger players should begin to show more corporate swagger.

Conservation imagery tracked this transition, with organizations turning to advertising agencies (which sometimes donated their time) for input into their public appearance. These interventions have not always been popular, particularly with the old-school membership. In 1986, for example, when the WWF invited Landor Associates of San Francisco to rationalize the several panda variants then in use, the brand consultants were critical of a slightly streamlined incarnation that Scott had approved. Landor judged that its nose was "too soft", its legs "too bow legged", its mouth "one-sided" and overall it just looked a bit "old, sick, depressed"[5]. Scott, it is said, was mortified by the disappearance of his anatomically faithful, playful panda. But the new abstracted panda suggested by Landor was better suited to life in the rich, digital environment that was on the horizon.

> ***"It was inevitable that the bigger players should begin to show more corporate swagger."***

Through the stylized output of these advertising agencies, conservation organizations were also able to move quietly away from the narrow species-specific remit that a fine-art logo implied. Counterintuitively, the simpler and more abstract the design, the better able it seemed to communicate the increasingly complex business of conservation.

The case of the Wildlife Conservation Society (WCS), which currently manages the Bronx Zoo, several other wildlife attractions in New York and more than 500 projects in over 60 countries, offers one of the most punctuated examples of logo evolution. In 2001, also with assistance from Landor, the WCS dropped its animal brand for an image that reflected the breadth of its twenty-first-century interests: out went a pair of excitable antelopes — its 'leaping lopes' — and in came a dappled green square that could be either a forest canopy or coral reef.

It was during this phase of abstraction that we first got a glimpse of a human element in logos. Conservation International began life in 1987, represented by a rainforest with a hut tucked away beneath the canopy, emphasizing the multi-species complexity of ecosystems (Fig. 1b). Also in the late 1980s, the UK-based Marine Conservation Society ditched a finicky hermit crab in favour of a dolphin and human diving together, a logo that was updated in 2000 (Fig. 1c). In the 1990s, Friends of the Earth (FOE) International came up with an avian squiggle superimposed on a circle that could be either the Sun or a human hand (Fig. 1l). These are some of the earliest graphic acknowledgements of something that all conservation organizations now understand: humans are part of both the problems and the solutions. The BTO's new logo, launched last year, is a recent example of this human dimension to conservation artwork: the head of a generic bird replaced Peter Scott's gannet and doubles as the pupil of a human eye.

If there is one drawback to this shift away from depicting charismatic creatures, it is that an organization is probably less likely to catch the eye of a mass audience for whom animals still have perennial appeal. But this might be a price worth paying if the substitute makes the right impression on the governments, agencies, foundations and corporations that have become a major source of conservation funding.

B: CONSERVATION INTERNATIONAL; C: MARINE CONSERVATION SOC.; E: WWT.ORG.UK; F: THE CHASE

G: BTO; H: IUCN; I: RSPB; J: THE NATURE CONSERVANCY

The IUCN, for example, has never sought a fluffy image. This is partly because that public appeal was provided by the WWF, which was founded to raise funds for the cash-strapped union. It is also because a hefty chunk of the IUCN's revenue is clinched in diplomatic dealings, which might explain why, after the extinction of the 'flaming artichoke', the union was content to spend several decades tinkering with various permutations of its acronym. Finally, in 2008, the IUCN (with advice this time from New York advertising agency Young and Rubicam) settled for a blue ring encircling the organization's initials. "The blue 'C' of the logo represents the planet and the union," says John Kidd, the IUCN's head of global communications. "The IUCN works on complex issues, often with complex solutions, but the logo is clean, clear, simple and, over time, hopefully memorable."

GLOBAL APPEAL

With the emergence of a truly global culture, and global concerns such as acid rain, nuclear fallout and climate change, it makes sense that this kind of holistic, planetary design has become more common. Most of the Friends of the Earth network swapped the charity's abstract sun and hand for a bright green, marker-pen circle in 2001. "It is a very simple design, and the idea was to represent the Earth, sustainability, cycles and unity," says Ann Doherty, communications coordinator at FOE International.

Similarly, in 2007, the international environmental organization the Nature Conservancy wrapped its trademark oak leaves around a green sphere (Fig. 1j). "As we've expanded outside the United States, now to more than 30 countries, we've incorporated the round, globe-like symbol to represent our focus on protecting lands and waters around the world," says Valerie Dorian, director of brand marketing and strategic partnerships. Even more nationally focused outfits, such as the United Kingdom's Woodland Trust, have adopted circular or spherical designs that give a nod to the scale of the problem (see Fig. 1f).

Many of these trends — the abstraction, a human presence, the appearance of a global element — have come together in Conservation International's new brand. Part of the reason for axing the long-standing rainforest logo was that it did not reproduce well in miniature, a quality essential in today's relentlessly digital world. It also failed to reflect the breadth of the organization's twenty-first-century mission "to empower societies to responsibly and sustainably care for nature for the well-being of humanity"[6]. In other words, Conservation International is about more than just rainforests.

The logo was the result of a consultation with New York design agency Chermayeff & Geysmar. "What Conservation International needed was not a literal picture that illustrates every single area of their activities, but rather a new, suggestive, and potentially expansive mark," says Sagi Haviv, the agency's principal designer. His solution — a blue circle underlined in green — is supposed to represent "our blue planet, emphasized, supported and sustained"; it also evokes an abstract human figure into the bargain.

The WWF's symbol is the most obvious exception that proves this trend towards global imagery. The organization never got type-cast in a species specific role; this is probably because the Chinese Cultural Revolution prevented the WWF from becoming involved with pandas until 1980. By then its panda had become established as a symbol with a truly global appeal.

So what should we make of a journey that began with literal, fine-art creations and has reached abstract images that make only a passing reference to nature? The answer, like the logos we're left with, is pretty simple. Conservation is no longer just about a single species on the brink of extinction, the habitat it's found in or some wider ecosystem. Now it's about the future of the planet. That, of course, means it's really all about us. ■ SEE BOOK REVIEW P. 290

Henry Nicholls *is a science writer based in London. His latest book is* The Way Of The Panda *(Profile, 2010).*
e-mail: henry@henrynicholls.com

1. Bonner, R. *At the Hand of Man: Peril and Hope for Africa's Wildlife* (Simon & Schuster, 1993).
2. 'This is the symbol of the World Wildlife Fund.' Max Nicholson Archive, Linnean Society of London, EMN 8/7.
3. *Oryx* **24,** 118–119 (1990).
4. Hughes-Evans, D. & Aldrich, J. L. *The Environmentalist* **1,** 91–93 (1981).
5. Schwarzenbach, A. *Saving the World's Wildlife: WWF's First Fifty Years* (Profile Books, 2011).
6. New Logo for a New Mission — Conservation International. (2010); available at http://go.nature.com/us6yap

CORRECTION

In the Comment article 'NASA: what now?' (*Nature* **472,** 27–29; 2011), the picture of the space shuttle *Challenger* disaster in 1986 was wrongly identified as that of *Columbia* in 2003.

Iconic species such as the African elephant have been a focus of conservation efforts for many decades.

M. HARVEY/WWF-CANON

CONSERVATION

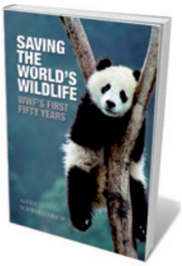
Bringing back wildlife

A scholarly history of the WWF highlights the gap between advocacy and action, finds **Nancy Lee Nash**.

The conservation group WWF is now half a century old, and *Saving the World's Wildlife* documents that history. This is a splendidly researched and written work. But whoever came up with the title should be spanked and sent to sit in the nearest natural history library for as long as it takes to list every known species that has gone extinct or become seriously threatened since 1961.

Alexis Schwarzenbach's history is more about an organization than Earth's misery. Even the most optimistic current reports of the planet's wildlife survival status — including every species from big mammals to birds, amphibians, fishes, trees, ferns, fungi and microorganisms — suggest there is little to celebrate. WWF International currently claims the largest membership, some 5 million, of any environmental organization worldwide. (Hong Kong, by comparison, has a population of 8 million.) Yet the WWF's success as a brand has eclipsed its original remit to protect species.

At its birth 50 years ago, the WWF was welcomed as a significant platform for conservation. Wildlife was both literally and metaphorically its middle name. It was created as a fund-raising organization to support the valuable scientific research of the International Union for Conservation of Nature (IUCN), which was founded at a conference of the United Nations Educational, Scientific and Cultural Organization in 1948. A group of UK-based visionaries who were passionately focused on the urgent importance of saving endangered animals and plants set up the WWF. Meeting these signatory heroes in the book and learning about their pioneering work is inspiring.



Saving the World's Wildlife: WWF's First Fifty Years
ALEXIS SCHWARZENBACH
Profile Books: 2011. 352 pp. £25/$39.95

Keen newcomers with a similar heartfelt mission arrived in later years, but the WWF continued in elder statesman mode — not least because of an impressive roll-call of royal stalwarts. Prince Bernhard of the Netherlands and Britain's Prince Philip were international presidents for many years each, and other royals were national presidents, adding allure, and commendable and significant support. The organization and its movers and shakers became widely known.

The WWF could have hired an agency to produce a glossy anniversary publication. Instead, to their credit, they commissioned an eminent Swiss historian. Schwarzenbach, who has previously written about Albert Einstein, studied WWF's scrupulously kept records and interviewed numerous living associates. None of them, he claims, tried to influence his interpretation of the WWF's history. It is all as precise as a Swiss watch.

Schwarzenbach faithfully records as much information about personalities, places and projects as his mission requires and space allows. Detailed notes, graphs and milestones are added in 30 pages of small print. Yet, surprisingly, these reveal that for almost half of those years there was hardly a pebble of significance. The period from 2004 to 2007 is empty of even a sand grain; 2008 is notable because "certified sustainable palm oil enters the market". Yet during the same years, most serious ecologists were focused on a larger issue: vast monoculture as one of the worst enemies of wildlife biodiversity.

The mid-1980s saw a major WWF International marketing

➲ NATURE.COM
For more on saving the panda, see:
go.nature.com/3dk65d

makeover. It was allowed to trash co-founder Peter Scott's brilliant original panda logo. Anyone familiar with the living creature, as I am, finds that the WWF's present symbol looks more like a large black-and-white plastic dog than the panda in Scott's iconic rendering.

The 'Wildlife' in the title was also exchanged. 'World Wide Fund for Nature', a less evocative phrase, was now employed by WWF International headquarters and 29 of the 31 national organizations around the world. Steadfast WWF Canada and US refused to give up the original name. Knowing how strenuously these changes were fought at the time, it is surprising that the word 'wildlife' is resurrected in the title of this anniversary volume.

Many activists have found efforts in the past 25 years to be less a task of saving endangered animals than of employing them as 'poster critters' in yet another corner of modern business. Dispassionate as it is thorough, the book reveals some strange marketing strategies. Who would have thought the art of selling cigarettes internationally could morph into tactics to raise funds for conservation? Conservation is an expanding market and there are more than enough crises to go around.

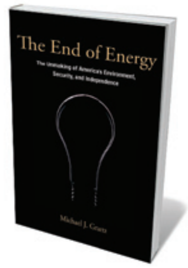
The author generously mentions other respected groups from the distant past. He highlights important forebears of the WWF, including US bodies such as the Wildlife Conservation Society, established in 1895; the UK Royal Society for the Protection of Birds, set up in 1889; and many others. But negligible mention of Harold Coolidge, IUCN founder and 'father of wildlife conservation', will be regarded as a serious oversight by many.

All conservation groups rely on public and governmental support. Readers and historians in the future will have to decide whether their strategies for acquiring funding are being matched by their efficiency in conservation. Perhaps they will interview the men, for the most part, who appear in this book, shown in suits (safari, business and formal); at home at royal events and in corporate boardrooms; and in some cases like deer caught in headlights.

Rejuvenation of the planet's diversity cannot depend only on organizations, their marketing teams, and a few heroes. It rests on the shoulders of each of Earth's leading large-brained primates. Thoughtful ones wish us luck. ■ SEE COMMENT P. 287

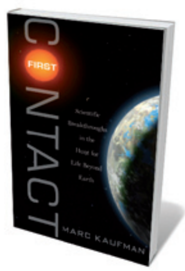
Nancy Lee Nash, *WWF consultant 1979–84, wrote the proposal for China–WWF contact, the first foreign involvement in protection of the giant panda; was instrumental in establishing WWF Hong Kong; and is a Rolex Award and UNEP Global 500 Laureate for her project Buddhist Protection of Nature.*

Books in brief



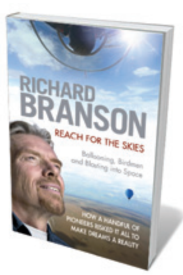
The End of Energy: The Unmaking of America's Environment, Security, and Independence
Michael J. Graetz MIT Press *384 pp. £22.95/$29.95 (2011)*
Energy policy has failed, according to law professor Michael Graetz. Pulling no punches, he argues that we must overcome the delusion we've lived with for 40 years that power is plentiful and cheap. By accepting inappropriate pricing of gas and oil, and being unwilling to negotiate around trade embargoes and for long-term solutions such as cap and trade, he argues that politicians have been complicit in distorting the energy market since the 1970s, with consequences for us all.

Mikhail Gorbachev: Prophet of Change: From the Cold War to a Sustainable World
Mikhail Gorbachev Clairview Books *344 pp. £14.99 (2011)*
As he turns 80, former Soviet president Mikhail Gorbachev publishes a selection of his writings and speeches on policy, philosophy and the environment. Since steering the cold war thaw of the 1980s, he has dedicated himself to supporting global peace, security and social justice as head of a non-governmental organization, Green Cross International. Having put sustainability at the centre of his cause, he has garnered the respect of many environmentalists and international leaders, who add tributes in this volume.



First Contact: Scientific Breakthroughs in the Hunt for Life Beyond Earth
Marc Kaufman Simon & Schuster *224 pp. $26.00 (2011)*
Venturing from the mouth of an Alaskan volcano to the bone-dry Atacama desert in South America, to a hot South African mine, science writer Marc Kaufman gets to grips with the young field of astrobiology. He describes how microbes survive in the harshest conditions on Earth, suggesting that organisms might cling on in extraterrestrial places like Mars. And he interviews researchers he sees as adventurers as well as thinkers: part Carl Sagan, part Indiana Jones, part Watson and Crick.

The Techno-Human Condition
Braden R. Allenby & Daniel Sarewitz MIT Press *192 pp. £18.95/$24.95 (2011)*
In this extended essay, ethicist and engineer Braden Allenby and science-policy expert Daniel Sarewitz explore what it means to be human in an age of technological extensions. From performance-boosting hormones and cognitive enhancements to re-engineered immune systems, our bodies and minds are being pushed into new evolutionary territories. They argue that terms such as human, technological and natural are outmoded, and should be rethought within a new techno-human relationship.



Reach for the Skies: Ballooning, Birdmen and Blasting into Space
Richard Branson Virgin Books *352 pp. £18 (2010/2011)*
Intrepid balloonist, entrepreneur and founder of the Virgin Group, Richard Branson shares his love of aviation in his latest book. Charting the exploits of daredevil pilots and aerial technologists through the ages, from the lavish balloons of the Mongolfier brothers to the many record-breaking feats of his friend the late Steve Fossett, he heralds the risk-takers who have made everyday flying possible. He looks ahead to commercial space travel, which he hopes to establish through his own company, Virgin Galactic.

M. ACHAKZAI/EPA/CORBIS

Epidemics can arise in the wake of extreme weather events such as the 2010 floods in Pakistan.

CLIMATE CHANGE

A dose of reality

Tony McMichael welcomes an overview of an under-recognized climate issue: the threat to human health.

Climate change may be an overarching issue of our times, yet most public debate about its consequences borders on the myopic. The long view — that human-induced global warming endangers Earth's basic life-support systems and, ultimately, people's well-being, health and survival — is obscured by more immediate preoccupations with its effects on electricity prices, gross domestic product and iconic ecosystems such as coral reefs and Arctic wildlife.

In *Changing Planet, Changing Health*, physician and public-health expert Paul Epstein and science journalist Dan Ferber explore far-reaching issues around climate and human health. The book covers the direct and indirect impacts of climate change on infectious diseases, food yields and nutrition, freshwater flows, air quality and the health consequences of community hardship and lost livelihoods. Epstein is a pioneer in this field, one of a handful of medical graduates who, in the early 1990s, focused on the emerging links between climate change and human health. Now associate director of Harvard Medical School's Center for Health and the Global Environment, he feeds into the book decades of field work, learning, teaching, research and advocacy.

Epstein opens with a personalized account of his grounding in environment and health — a semi-autobiographical leavening that recurs in several later chapters. In 1976, working in public health in post-independence Mozambique, Epstein quickly saw how the interplay of environmental, social, economic and political factors determines patterns of disease in communities. His recognition of the ecological basis of health problems, particularly infectious diseases, led him to apply the ideas of systems theory to this area, including the risks of infection in animals and plants in disturbed ecosystems.

In an excellent exposition, he discusses environmental microbiologist Rita Colwell's discovery in the 1990s of how algal blooms in warm waters nurture cholera bacteria, which then enter the marine food chain and, eventually, human consumers. This idea of an aquatic ecosystem as a culture medium for cholera bacteria, hosted by algae that proliferate in waters enriched by run-off nutrients and warming, has extended our understanding of cholera transmission well beyond the mechanistic person-to-person spread by faecal contamination.

The authors provide well-chosen historical illustrations of the social, economic and political triggers of infectious-disease outbreaks. Surprisingly, they do not mention other evidence for specific climatic influences on disease. The book attributes epidemics of bubonic plague in Europe in the sixth and fourteenth centuries and typhus in Silesia in the nineteenth century to combinations of political decline, social unrest, economic exploitation and poverty. Yet there is evidence that many such epidemics arose in the wake of regional climatic fluctuations. For example, major outbreaks of bubonic plague followed weather-driven increases in the number and activity of rodent hosts and flea vectors. Extremes of cold and wet may have increased the susceptibility of Silesian workers to typhus.

Changing Planet, Changing Health: How the Climate Crisis Threatens Our Health and What We Can Do About It
PAUL R. EPSTEIN AND DAN FERBER
University of California Press: 2011. 368 pp. £18.50/$29.95

The book's main focus is the risk to human health from current and future climate change. The expectation, corroborated by some emerging evidence, is that human-induced climate change will expose populations to more intense floods, storms and fires; more severe heatwaves; and changes in the rates and ranges of infectious diseases, which will encroach or increase in many regions, yet decrease in others. It is also likely to impair global crop yields (but with regional variability), and disrupt societies and their health through displacement, regional setbacks in farm yields and rural family incomes.

The authors discuss the still-debated relationship between climate change and malaria; how future climate-change scenarios may affect patterns of dengue fever, Rift Valley fever and other infectious diseases; and the risks of respiratory diseases from climate-related increases in air pollution.

They also examine the adverse effects of climate on food supplies, nutrition and child survival, particularly in populations that have low food security; as well as the detrimental impacts of ocean warming, increased acidity and other environmental shifts on marine ecosystems, fish harvests and toxins in seafood. They consider climate effects on forest ecosystems and biodiversity, such as increased beetle infestations in US pine forests and the subsequent forest-fire risks of dead trees, and the susceptibility of poor and vulnerable communities to injury, death and disease from extreme

➲ NATURE.COM
For more on malaria control, see:
go.nature.com/p8rhen

weather events. Epstein and Ferber conclude by exploring solutions. In their own guardedly optimistic words, they aim to show "how we must transform the way we power society and organize our economy to preserve a liveable planet. But it is also about the incredible opportunities that will arise once we do."

They examine the political and cultural impediments to climate-change acceptance and policy response in the United States. Technological options for achieving a low-carbon economy are detailed, including the relative risks and benefits of alternative energy sources. The biggest challenge, they say, will be to modify consumption behaviours and expectations. Yet many benefits beckon as societies undertake transformative technological and behavioural actions to abate climate change. Localized health gains would result — such as breathing cleaner air if fewer fossil fuels were burnt, and having healthier bodies if we relied less on cars.

They boldly explore possibilities for transforming the global economic system, including the creation of a new international financial architecture to replace the current outmoded system. Inequalities in power and profit in current worldwide trade need to be remedied, and greater weight must be given to achieving international equity, population well-being and environmental sustainability. Although the authors recognize the merits of market-based economies, they argue that the prevailing model is inadequately regulated and increasingly damaging in a world where commodities and capital flow across national borders, often to the disadvantage of poorer countries.

The authors emphasize that the great challenge for the world community — given the climate threat to the future of our species and to those we depend on — is to build and benefit from a sustainable set of social values, economic priorities, low-impact technologies and coordinated national and international institutions. We urgently need to comprehend that the risks to human health and populations will be great if climate change continues on its current trajectory. Because human health is 'the bottom line' at which the many adverse consequences of climate change will converge, *Changing Planet, Changing Health* is an excellent corrective for climate-change myopia. ■

Tony McMichael *is a professor at the National Centre for Epidemiology and Population Health at the Australian National University, Canberra, Australia.*
e-mail: tony.mcmichael@anu.edu.au

ART

Grime and punishment

Daniel Cressey finds there is more to dirt than disease at a London exhibition.

Of more than a dozen exhibitions held at London's Wellcome Collection in the four years since it opened, curator Kate Forde admits that this is the first on which she has consulted the gallery's cleaners.

Dirt attempts to make visitors think again about mundane muck. The exhibition takes inspiration from the ideas of anthropologist Mary Douglas, who noted that what we consider to be dirt is based on context — a theme that is explored in six cities around the world and across human history. Accompanying events include conversations with a street sweeper and with the wonderfully titled 'chief flusher' of local utility firm Thames Water.

Starting in the Netherlands in the seventeenth century, the displayed pottery, tiles and engravings show the 'cleanliness is next to godliness' obsession of the Delft Christians in their spotless homes. We encounter favourites of science historians: one of Antonie van Leeuwenhoek's surviving microscopes sits alongside a copy of Robert Hooke's *Micrographia* with its drawings of microbes. Van Leeuwenhoek's description of the tiny 'animalcules' found in dirty water revealed the creatures previously hidden in the grime.

The health theme continues to unfold through the squalor of Victorian times, with

Dirtiness can have social implications.

MUSEUM OF LOCNON/WELLCOME IMAGES

Dirt: The Filthy Reality of Everyday Life
The Wellcome Collection, London. Until 31 August 2011.

John Snow's famous map on the spread of cholera in London, and medical apparatus from Joseph Lister's pioneering use of antiseptics in Glasgow. But things soon take a darker turn as the show moves on to reveal how ideas of cleanliness were co-opted for the oppression of those who were considered dirty.

A 1911 exhibition on hygiene in Dresden, Germany, is the starting point for the fourth cityscape. The chilling quote from the country's Law for the Prevention of Genetically Diseased Offspring from the 1930s hints at where this cleanliness obsession is headed: "Eradicate biologically inferior hereditary taints. Promote a gradual cleansing of the nation's ethnic body."

The exploration of the use of the word 'dirty' continues with the section on the people who clean the latrines of New Delhi and Kolkata, and their lowly place as 'untouchables' in India's caste system. Here is placed the most deliberately provocative of the modern art pieces displayed alongside the historical works throughout the show: several large grey slabs that make up the 'Anthropometric Modules' constructed from human faeces, by artist Santiago Sierra.

Most of the artworks are blunt in carrying Douglas's message that dirt is not bad, simply contextual. Angela Palmer's indistinguishable air samples, collected in 2007 from the cleanest and most polluted sites in the world, make invisible again what van Leeuwenhoek strove to make visible. Serena Korda's bricks — made with dust collected from locations ranging from J. G. Ballard's books to the UK House of Commons debating chamber — transform the dirt by giving it utility.

The most powerful proof that dirt is 'matter out of place' and can be redeemed by repositioning it is found at the end of the exhibition. Here the show moves forward in time, to New York's Staten Island in 2030. By then, it is hoped that Fresh Kills, one of the world's largest waste-dumping sites, will have been transformed into an urban park.

With exhibits such as a carpet made of dust, *Dirt* also generates what it celebrates. So it is a good job that the gallery's cleaners are on board. ■

Daniel Cressey *is a reporter at* Nature.

J. SHAPIRO/PLANET GREEN

Actor Isabella Rossellini takes on the role of Charles Darwin in her film *Animals Distract Me*.

Q&A Isabella Rossellini

Animal distractions

Isabella Rossellini, star of films including Blue Velvet *(1986) and* Big Night *(1996), has made a series of short films on the mating rituals of insects and sea creatures. As her latest humorous biopic debuts in the United States, Rossellini explains why she is fascinated by animals.*

Why have you made films about animals?
I have wanted to make funny films about animal behaviour since I was a teenager. The first short films that I made on this topic were a web series called *Green Porno* for the Sundance Channel in 2008. I did 18 episodes about how animals reproduce. Then I did ten episodes called *Seduce Me,* about all the different courtship strategies, from birdsong to colourful plumage to rape. And then the Discovery Channel commissioned a one-hour special, *Animals Distract Me*, which had its premiere at the Sundance Film Festival this year and will be broadcast in the United States on 23 April.

What's in the new film?
The two-minute format of my previous films didn't work for the Discovery Channel. I had to think of an hour-long story, but I wanted to have little episodes to convey curiosities about animals. The device we chose was a day in my life, where I get up, go to work as a model, have lunch in an elegant restaurant, study my script and go to bed. But I'm constantly distracted by animals, whether it is a pigeon in the street, a cockroach on my table, the meat on the menu or the guide dog that I'm raising. It is the way it has always been with me.

You also play Charles Darwin, in a series of surreal asides on evolution. Why?
When I started getting older and working less, I decided to go back to university and study animal behaviour. I'm still finishing my undergraduate degree in the history of science and environmental studies at New York University. I took several courses on Darwin. He was one of the first scientists to understand the importance of photography in scientific research. One of his books that fascinated me is *The Expression of the Emotions in Man and Animals* (1872). Darwin asked himself the same question an actor asks: why are some gestures understood internationally, others only in one culture? He imagined that squinting began as a way to reduce glare and then became attached to the process of thinking. He had ideas that were fascinating; almost naive, but genius too.

Did the film have a scientific adviser?
Yes, John Bohannon, a visiting researcher at Harvard University in Cambridge, Massachusetts, who writes for *Science*. When I tried to work with other scientists, it was too complicated. Film has to be visual, fast and simple. John makes sure I don't use a word that is too mistaken. I want to be scientifically correct, but sometimes there are nuances that cannot be used. When I did a little poem for *Green Porno* about animals that change sex, I used the word transsexual. It wasn't the scientific term, but John thought it was okay.

How have your films been received among scientists?
I thought somebody was going to scold me for *Green Porno*. But it has been quite well received by the scientific community. It is now part of the *Sexual Nature* exhibition at London's Natural History Museum. John organized a fan club at Harvard, a group of eight scientists who came to the Toronto Film Festival dressed in homemade hats shaped like animal penises. Comedy hasn't been used much in films about biology or the environment. It comes naturally to me. Animals make me laugh. I wanted to capture not only the wonder of nature, the awe and intimidation, but the joy and humour too.

Why use humour to convey science?
Everybody knows that some animals are hermaphrodites, but it is said in such heavy terms that people assume it is boring. The gloom and doom of the environmental message, which has worked well for fund raising, has conquered the maximum audience that can be reached with that language. Now there is a wall. Maybe if we can talk with humour and lightness, we can recruit more people to listen to the variety of nature.

Is it a challenge to convey the complexities of science on screen?
When theorists debate, the public is cut out because they don't know the shorthand. I recently read a book by E. O. Wilson. He has to invent words to explain what he has observed in ants, but I cannot use the word superorganism in a film. It becomes complicated unless I find something physical that can explain it visually. When I'm studying something complex, I draw. That's how my films come about. The scripts are a series of drawings.

What's your next project?
I would like to make a film about motherhood. I'm about to do an independent study on a series of women biologists who looked at maternal instinct, hoping that out of that will come a little film. I'm still looking for financing. If I was younger, I would have liked to make field documentaries about animals. But they are already excellently made, and at my age I can't imagine doing it better than David Attenborough. So I'll tell you a story the way I would like to tell it. ■

INTERVIEW BY JASCHA HOFFMAN

Animals Distract Me
ISABELLA ROSSELLINI
On the Discovery Channel at 10 p.m. (EDT) on 23 April.

CORRESPONDENCE

Doing the maths on the green economy

In their 21 February report *Towards a Green Economy*, the United Nations Environment Programme (UNEP) claims that "a green economy grows faster than a [conventional] brown economy, while maintaining and restoring natural capital". We contend that this claim is founded on flawed assumptions.

First, it is based on an inadequate target for reducing energy-related carbon dioxide emissions. The reduction by 2050 under UNEP's G2 green-economy scenario is only 35% relative to 2011 emissions. This cannot achieve the 450 parts per million (p.p.m.) target arising from IPCC's Fourth Assessment, let alone more recent calls for 350 p.p.m. Also, UNEP is assuming that the investment needed to achieve this target is additional to (rather than just a reallocation of) business-as-usual investment funds, ignoring crucial questions about financing. This additional 'green' investment is 10% higher than that applied to the 'brown' scenario, causing partiality in the analysis.

In addition, the claim hinges on the modelling of the world as a single economy, which overlooks substantial regional inequities and obscures the interplay between economic growth, social justice and climate-change mitigation. Assuming no change in the 2007 15:1 ratio in GDP per capita between high- and low/medium-income countries, for example, UNEP's G2 scenario requires a rate of decline in average CO_2 intensities in both country income groups of 3.9% per year. By 2050, this would lead to an increase in absolute per capita income differences of US$104,000 from $42,000 per capita in 2011, and would reduce emissions by only a third.

For full convergence of living standards by 2050, coupled with the 80% reduction in global CO_2 emissions required to meet a 450 p.p.m. target, we would need a decline in average CO_2 intensities in both regions of 8.2% per year.

Taking these factors into account, we cannot presume that a green economy would grow faster than a brown one.

Peter Victor *York University, Toronto, Canada.*
peter@pvictor.com
Tim Jackson *Centre for Environmental Strategy, University of Surrey, Guildford, UK.*

Academia resists clean-up in Romania

A law to reform Romania's education system, approved in February (*Nature* **469,** 142–143; 2011), is already floundering. University staff seem willing to compromise their academic standards when it comes to political matters.

The law bans members of parliament from holding positions of power in universities. But even as the government is busy defining the law's regulations, parliamentarian university rectors from across all the political parties are insisting that the law should not apply to them because they were elected to parliament before the legislation came into effect.

Legal committees of both houses of the Romanian parliament decreed in February that there should be no incompatibility between academic functions and parliamentary positions. But the ministry of education intends to ask the National Integrity Agency, which controls corruption and mismanagement in government, to investigate. Rectors enjoy extra financial benefits when they become parliamentarians, and have the power to maintain the status quo in Romania's inbred, nepotistic and underperforming university system.

To disrupt this system, the next generation of academics needs to be out-bred. External reviewers from academically advanced European Union countries should play a part in overseeing promotions and funding in Romanian universities.

Liviu Giosan *Woods Hole Oceanographic Institution, Massachusetts, USA.*
lgiosan@whoi.edu

Science failed to self-regulate

As initiators of the open letter that contributed to the resignation of the German defence minister Karl-Theodor zu Guttenberg after allegations of scientific misconduct (*Nature* **471,** 135–136; 2011), we find the reaction of the scientific community disappointing.

Only a handful of senior scientists challenged Chancellor Angela Merkel's point that she had hired a defence minister, not a research assistant: they criticized her for conveying the impression that the issue, and thus academic integrity in general, was politically irrelevant. Some of Germany's most important research institutions took their time to issue cautious statements against the dismissal of fundamental academic principles. A few notable organizations have remained silent.

The affair also raises questions about the relationship between academia and society. Why has it been so hard to convey the importance of academic integrity, given that politicians in Germany and elsewhere regularly proclaim the advantages of the 'knowledge society'? And why did the professional self-regulation of science fail so miserably in the Guttenberg case? This case should mark the beginning, not the end, of a serious debate on academic rigour in Germany.

Frederik Trettin, Tobias Bunde *University of Konstanz, Germany.*
frederik.trettin@uni-konstanz.de

Fund biodiversity collections

Universities are crucial for holding specimen collections in megabiodiverse, developing regions such as Brazil (see *Nature* **471,** 164–165; 2011). Only local institutions can gather information with the necessary resolution to answer biological questions at a regional scale. But such collections suffer from serious funding shortages, mostly because they are not included in a permanent university budget.

A huge volume of material is already stored in university collections throughout Brazil, and these are proliferating as higher education expands in the country. Universities with scientific collections provide a rich environment for recruiting and training evolutionary biologists who will ultimately deal with pressing biodiversity issues. Institutional support would guarantee the collections' continuation beyond the working life of their curators.

The University of São Paulo, Ribeirão Preto, is now uniting its scattered collections into a biodiversity museum — an initiative sparked by a grant of almost US$800,000 this year from FAPESP, São Paulo state's research funding agency. Other universities could follow this lead in coordinating biodiversity research to help meet conservation and management priorities.

Flávio A. Bockmann *and eight co-authors*, University of São Paulo, Ribeirão Preto, Brazil.*
fabockmann@ffclrp.usp.br
**A full list of authors is online at http://dx.doi.org/10.1038/471295d.*

Jürg Tschopp

(1951–2011)

Immunologist whose discoveries transformed patients' lives.

Jürg Tschopp stood apart from most of his colleagues in immunology in that his discoveries in fundamental research brought striking and almost immediate benefits to patients suffering from painful, debilitating diseases. Of his numerous findings, one led to a treatment for the autoimmune disorder systemic lupus erythematosus, and another brought dramatic relief to patients with gout.

Tschopp, who died on 22 March aged 59, trained as a biophysicist and obtained a PhD from the University of Basel in Switzerland in 1979. Yet it was immunological questions that captured his attention. He did postdoctoral work with molecular immunologist Hans Müller-Eberhard at the Scripps Clinic in La Jolla, California, before joining the Department of Biochemistry at the University of Lausanne in Switzerland in 1982.

An early milestone in his career was the characterization of the final step in a signalling pathway called the complement cascade. This is a crucial component of the innate immune response, which protects animals from an infection through a generic mechanism, rather than on the basis of previous 'remembered' encounters with an infectious agent.

In the early 1980s, immunologists knew that one protein involved in the complement cascade, C9, could kill bacteria, but how it worked was a mystery. Tschopp showed that C9 essentially drills holes in the cell wall of a bacterium by assembling into a pore within its cell membrane. Vital macromolecules leak out through the pore, and the bacterium quickly dies.

Subsequently, Tschopp found that a class of white blood cells called T cells destroy tumour cells or those harbouring viruses using a similar mechanism. In this case, these T cells release perforin, a channel-forming protein that inserts itself into the membrane of the target cell, allowing an enzyme called granzyme to enter the cell and destroy it.

Tschopp relished turning up clues in the detective stories unfolding in his laboratory. To his and everyone else's surprise, T cells in mice lacking the gene encoding perforin could still destroy target cells. The back-up pathway turned out to involve a membrane protein called the Fas ligand. When this protein binds to its receptor protein Fas on target cells, a molecular cascade begins that activates a cell-destroying enzyme similar to granzyme.

It was Tschopp's later discoveries that had a direct impact on patients. Along with other groups, he identified BLyS as a ligand for a receptor that drives the proliferation of B cells — a type of white blood cell that produces antibodies. These antibodies are proteins that recognize and neutralize certain types of infecting agents as part of a longer-term defence strategy.

J. STÜCKER

Tschopp's discovery catalysed the development of a compound called Benlysta (belimumab) by the pharmaceutical company Human Genome Sciences in Rockville, Maryland. Benlysta neutralizes the BLyS ligand, and so can treat immune disorders in which B cells produce antibodies against the body's own healthy cells. Benlysta received approval from the US Food and Drug Administration just last month as a treatment for the chronic autoimmune disorder systemic lupus erythematosus. It is the first new lupus drug in more than 50 years.

About ten years ago, while working on BLyS, Tschopp mentioned to me over dinner that he wished to devote the rest of his career to discoveries that made a difference to patients' lives. Soon after, he began his quest to characterize the inflammasome signalling complex.

The inflammasome complex is a protein pathway that starts the production of two cytokine proteins that orchestrate the inflammatory response. In a scientific tour de force, Tschopp's laboratory showed that the tender, swollen joints and excruciating pain associated with gout are caused by uric-acid crystals activating the inflammasome. Based on these findings, Tschopp persuaded the clinical-rheumatology group at the University of Lausanne to conduct human trials on blockers of one of the cytokines for the treatment of gout. The results were stunning. Patients enjoyed almost immediate, dramatic and prolonged relief.

This success thrilled Tschopp. He took particular pleasure in showing the clinical data at meetings to support the claim that basic research yields significant clinical achievements. For the inflammasome discovery and other advances, he was showered with accolades and prizes, including the Louis-Jeantet Prize in 2008 and the Novartis Prize in Clinical Immunology in 2010.

Jürg was a formidable athlete and competed in the decathlon in his youth — a competition consisting of ten events ranging from long jump to discus. Throughout his life, he maintained a vigorous outdoor lifestyle, skiing and playing tennis with his family and jogging with his precious dogs. Tragically, he died while hiking in his beloved Swiss mountains.

He once told me of how he had swum in a lake with his retriever, realizing too late that his wallet had been in his pocket and had fallen out. Suddenly he spotted his dog running furiously towards him, wallet in mouth! The story is fitting because Jürg had an uncanny ability to dive into the murky waters of complex scientific issues and emerge, faster than the rest of us, with the answer.

As we mourn Jürg's unexpected death, we can take solace that his work will live on and benefit many patients. ■

Vishva Dixit *is vice-president of research at Genentech, South San Francisco, California 94080, USA.*
e-mail: dixit.vishva@gene.com

DEVELOPMENTAL BIOLOGY

Remarkable role for the placenta

The placenta does not only act as the essential link between the developing embryo and its mother. Unexpectedly, it is also a source of serotonin — a neurotransmitter that is crucial for embryonic brain development. SEE LETTER P.347

RON MCKAY

By the time the fertilized egg has undergone only a few divisions, it has already divided into the prospective embryo and the cells that will form the placenta, the embryo's crucial link to its mother. Because the embryo and placenta share the same genome, they can be seen as twins. The very early developmental events that form the placenta exert a continuing fascination, like any good creation story[1,2]. But on page 347 of this issue, Bonnin *et al.*[3] focus on events beyond these initial stages and show that the placental cells synthesize a signal — the neurotransmitter serotonin — that reaches the embryonic forebrain at a crucial stage of development and so may influence brain function in the adult mammal.

Serotonin is only small — it has just 10 carbon, 12 hydrogen, 2 nitrogen and 1 oxygen atoms. But it has powerful effects when in the right place at the right time. Receptors for serotonin are the targets of some of the most widely prescribed drugs for regulating mood. Although in the adult only a few neurons in the hindbrain make serotonin, their axonal processes project throughout the brain. Bonnin and colleagues' paper[3] opens with the observation that serotonin levels in the cortical region of the developing mouse brain are not affected when the production of this neurotransmitter in the hindbrain is blocked. So where does this signal come from? Candidate sources included some other part of the developing embryo, the mother and the placenta.

In the detective work that followed, the authors used many techniques, but their central experiment is easy to understand. They devised a new technology for regulating placental organ perfusion, and were thus able to directly test the biochemical functioning of the placenta. Serotonin (5-hydroxytryptamine) is synthesized through a single intermediate (5-hydroxytryptophan) from the amino acid tryptophan. Bonnin *et al.* find that increasing tryptophan levels in the maternal uterine artery — the major source of nutrients to the placenta — leads to elevated levels of serotonin in the fetal umbilical vein 15 minutes later (Fig. 1). This observation directly contradicts the prevailing idea that serotonin from the mother crosses into the embryo. The placenta itself makes serotonin and regulates its levels in the developing cortex.

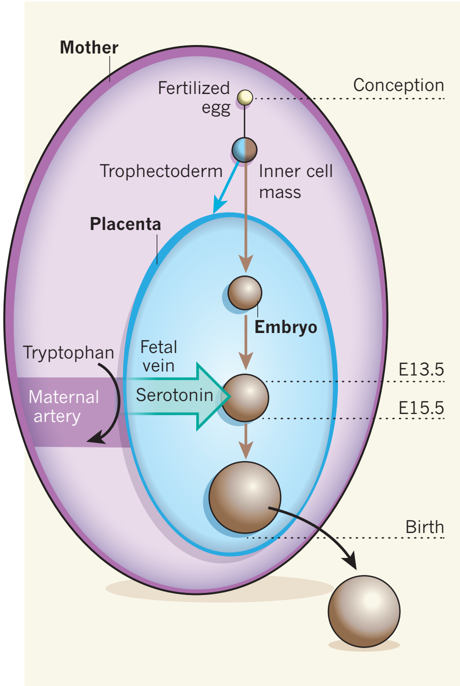


Figure 1 | Placental signals to the developing brain. Following conception, the fertilized egg undergoes just a few cell divisions before the resulting cells separate into two genetically identical populations: cells that will form the embryo (the inner cell mass) and cells that will form the placenta (trophectoderm). Bonnin *et al.*[3] show that, around embryonic days 13.5–15.5, placental cells synthesize serotonin from tryptophan delivered to the maternal uterine artery. They also show that serotonin synthesized by the placenta regulates the level of this neurotransmitter in the forebrain of the developing fetus. Because these events occur at a time when serotonin is known to regulate forebrain development, they may have a lasting influence after birth.

Bonnin and colleagues particularly focus on the developmental stage between embryonic days 13.5 and 15.5, when the major neuron types of the mouse cerebral cortex are generated. Although the researchers did not investigate the effect of serotonin on brain development, there is good evidence[4,5] that this neurotransmitter influences the numbers of cerebral-cortical neurons, their location and how they wire up to other parts of the brain. So the placenta seems to signal information that regulates early cortical development.

Now that the source of serotonin has been defined, scientists can address the regulatory and long-lasting effects of this signal. Bonnin *et al.* demonstrate that serotonin does not cross the placenta. So, are serotonin levels in the early embryonic cortex affected by other maternal modulators? In adults, specialized transporter proteins rapidly remove serotonin from the sites at which it is active. These transporters are targets of drugs that regulate mood, and naturally there is interest in the possibility that such drugs would influence embryonic brain development[6].

Does the placenta control other neurotransmitters or hormones that have long-term effects on the brain? One possible candidate is the neurotransmitter dopamine. The placenta expresses catechol-*O*-methyltransferase (COMT), an enzyme that inactivates dopamine and also generates 2-methoxyestradiol. Pregnant mice that lack COMT show symptoms of pre-eclampsia — a severe condition in humans that can affect many organs in pregnant women[7]. Notably, 2-methoxyestradiol corrects these symptoms, with this signal passing from the placenta to the mother. The effects of placental COMT on embryonic brain development are not known, but in the light of the current report[3] they are of great interest.

The placental cell that synthesizes serotonin is the syncytiotrophoblast — a giant cell derived from the trophectoderm, the precursor cells of the placenta. This cell also makes high levels of monoamine oxidase A (MAO-A), an enzyme that degrades monoamine neurotransmitters including serotonin, dopamine and noradrenaline. In humans, mutations in MAO-A lead to psychiatric and behavioural disorders that affect only males[8]. It will be interesting to determine whether placental MAO-A has a long-lasting effect on brain development.

The first placental cell types made by the embryo are those of the trophectoderm (Fig. 1). Work on early embryos and on

pluripotent cell lines that differentiate into trophectoderm reveals differences between various mammalian species[9,10]. However, even though the human and mouse placentas may not influence their embryos in the same way, it is likely that this quiet twin from our prenatal life affects our development more directly than was previously thought. ■

Ron McKay *is at the Lieber Institute for Brain Development, Baltimore, Maryland 21205, USA.*
e-mail: ronald.mckay@libd.org

1. Gardner, R. L. *J. Embryol. Exp. Morphol.* **88,** 303–326 (1985).
2. Johnson, M. *Reprod. Biomed. Online* doi:10.1016/j.rbmo.2011.03.001 (2011).
3. Bonnin, A. *et al. Nature* **472,** 347–350 (2011).
4. Vitalis, T. & Parnavelas, J. G. *Dev. Neurosci.* **25,** 245–256 (2003).
5. Bonnin, A., Torii, M., Wang, L., Rakic, P. & Levitt, P. *Nature Neurosci.* **10,** 588–597 (2007).
6. Frost, D. O., Gibb, R. & Kolb, B. *Neuropsychopharmacology* **35,** 344–345 (2010).
7. Shenoy, V., Kanasaki, K. & Kalluri, R. *Trends Endocrinol. Metab.* **21,** 529–536 (2010).
8. Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H. & van Oost, B. A. *Science* **262,** 578–580 (1993).
9. Rossant, J. *Nature* **471,** 457–458 (2011).
10. Kunath, T. *Cell Stem Cell* **8,** 241–242 (2011).

MATERIALS CHEMISTRY

Spot-on healing

Rubbery polymers have been made in which damage is healed by exposure to light. The healing mechanism allows localized, on-demand repair, and might help to extend the lifetimes of materials for many applications. SEE LETTER P.334

NANCY R. SOTTOS & JEFFREY S. MOORE

Although ultraviolet radiation can be harmful to humans, Burnworth *et al.*[1] report on page 334 that it can be harnessed to heal scratch damage in specially designed elastomeric polymers. The authors' optically healable materials consist of polymers that contain supramolecular metal complexes, which absorb ultraviolet light and convert it into heat, triggering the mending process. Because light can be focused precisely onto damaged spots, this approach could potentially be used to locally heal polymers bearing loads.

Tough, rubbery plastics have revolutionized consumer products ranging from simple bags and storage containers, to bladders and inner tubes in tyres, to protective coatings and linings. Yet the materials used in these applications are highly susceptible to damage caused by scratches, cuts and punctures, which result in loss of function and contents, contamination, safety hazards or just inconvenience. Damaged plastics are typically discarded in landfill, or in some cases repaired with a manually applied patch. Healable polymers offer an alternative to the damage-and-discard cycle and represent a first step in the development of polymeric materials that have much greater lifespans than currently available materials.

All classes of polymer could potentially be healed, and several successful strategies have emerged for storing and triggering healing functionality in a wide range of polymers and polymer composites[2]. In many cases, cracks or scratches can be repaired by heating the material above its glass transition temperature, then applying pressure to promote contact, wetting, diffusion and re-entanglement of the polymer chains[3]. But this brute-force method is slow and difficult to apply in practice.

A more intrinsic approach involves modifying polymer structures at the molecular level to include chemical groups that bond reversibly or dynamically, so that bonds broken when the material is damaged can simply be re-formed. One successful strategy for intrinsic healing is to synthesize polymers from components that transform reversibly from the monomeric state to a crosslinked polymeric state through the addition of external energy, most often heat[4]. Reversible gels, which switch between gel and viscous liquid states, are underpinned by this mechanism and have been under development for decades[5]. For structural polymers, a major step forward was the discovery of a mendable polymer[6] based on the reversible Diels–Alder reaction (in which two carbon–carbon bonds are formed). With the application of heat (120–150 °C) and modest pressure, cracks in these polymers can be repeatedly healed.

Another promising scheme for intrinsic healing relies on supramolecular assemblies. Unlike conventional polymers, in which the monomers are bound together covalently, supramolecular polymers consist of repeating units held together by non-covalent bonds. Such bonding is reversible, and so, if broken, the bonds can re-form within a certain limited period of time after the material has been damaged[7,8]. One example is a self-healing, rubber-like material that has been made from a hydrogen-bonded supramolecular network[9] — if broken ends of the damaged material are brought back together at room temperature, hydrogen bonds re-form between molecules in the ends, and the material heals. Other healable supramolecular elastomers have been reported[10], but these require a lot of heat to repair damaged areas.

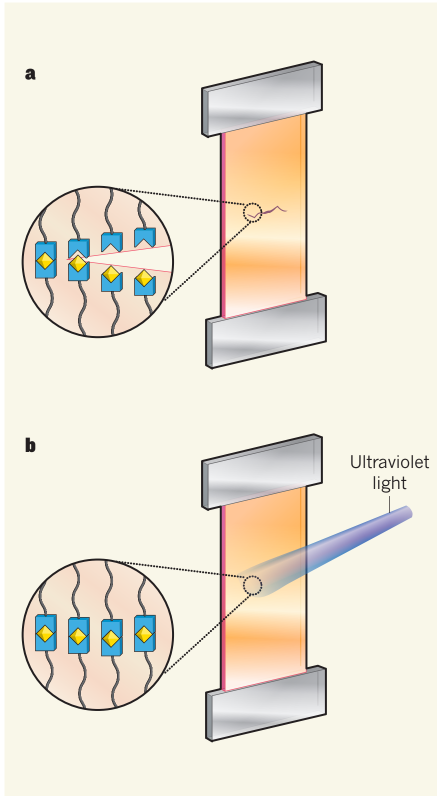


Figure 1 | Healing on demand. Burnworth *et al.*[1] have prepared polymeric materials in which scratch damage is healed on exposure to ultraviolet light. **a**, This cartoon shows a scratched thin film of the material under load. The inset depicts the molecular structure of the material, which consists of hydrocarbon chains (black) terminated with ligands (blue) bound to metal ions (yellow). The metal–ligand complexes are ruptured at the scratch. **b**, When ultraviolet light is shone on the scratch, it excites the metal–ligand complexes, which then relax to their ground states and generate heat. This heat causes the surrounding material to soften, so that the ruptured complexes re-form and the scratch is mended.

Burnworth *et al.*[1] now report that thermal mending of supramolecular polymers can be triggered by exposure to light. To construct optically healable materials, the authors used hydrocarbon oligomers — short polymer chains — whose ends terminate with ligands that can bind to metal ions. When the authors added zinc ions (Zn^{2+}) or lanthanum ions (La^{3+}) to a solution of the hydrocarbons, metal–ligand complexes formed in which each ion was bound to two (in the case of Zn^{2+}) or three (in the case of La^{3+}) oligomers, linking the oligomers together to form long-chain supramolecular polymers. The polymers thus had a periodic structure in which a hydrophobic segment (a hydrocarbon section)

followed a polar segment (a metal–ligand complex), repeated many times over.

In the solid state, the polymer chains underwent phase separation — they folded in such a way as to form some domains that were rich in hydrocarbon segments and others that were rich in metal–ligand complexes. The combination of long-chain polymers and phase separation gave rise to a physically crosslinked network, a molecular architecture that makes the material tough.

The molecular characteristics of Burnworth and colleagues' polymers were ideally set for localized healing driven by ultraviolet radiation. First, the polymers undergo heat-induced depolymerization, presumably because the equilibrium between polymerized and non-polymerized material favours the latter at higher temperatures. Second, the metal–ligand complexes absorb and are excited by ultraviolet light, with the energy involved being efficiently converted to heat as the excited complexes return to their ground states. So, when the authors shone ultraviolet light at damaged films of the polymers, photothermal healing occurred because the light was converted into heat, which then caused the materials to depolymerize, liquefy and seal up cuts in the film (Fig.1).

Although supramolecular approaches to intrinsic healing hold great promise, the synthesis of a strong, stiff polymer with autonomous healing properties remains a key research challenge. In current state-of-the-art supramolecular polymers, healing does not occur automatically, and energy in the form of light or heat must be supplied while the metal–ligand bonds are still able to re-form. Burnworth and colleagues' work does not specifically address these problems, but it does provide exciting new opportunities. Because different metal–ligand complexes absorb light at different wavelengths, it should be possible to tune the wavelength at which materials heal simply by changing the complexes. This in turn offers the prospect of smart, force-sensitive polymers that change colour when damaged, with the colour being matched to the wavelength of light that triggers healing; irradiation with that light would thus cause localized healing only at the damaged spots. ■

Nancy R. Sottos *is in the Department of Materials Science and Engineering, and* **Jeffrey S. Moore** *is in the Department of Chemistry, University of Illinois, Urbana, Illinois 61801, USA.*
e-mails: n-sottos@illinois.edu; jsmoore@illinois.edu

1. Burnworth, M. *et al. Nature* **472,** 334–337 (2011).
2. Blaiszik, B. J. *et al. Annu. Rev. Mater. Res.* **40,** 179–211 (2010).
3. Kim, Y. H. & Wool, R. P. *Macromolecules* **16,** 1115–1120 (1983).
4. Bergman, S. D. & Wudl, F. *J. Mater. Chem.* **18,** 41–62 (2008).
5. Schultz, R. K. & Myers, R. R. *Macromolecules* **2,** 281–285 (1969).
6. Chen, X. *et al. Science* **295,** 1698–1702 (2002).
7. Brunsveld, L., Folmer, B. J. B. & Meijer, E. W. *MRS Bull.* **25,** 49–53 (2000).
8. Serpe, M. J. & Craig, S. L. *Langmuir* **23,** 1626–1634 (2007).
9. Cordier, P., Tournilhac, F., Soulié-Ziakovic, C. & Leibler, L. *Nature* **451,** 977–980 (2008).
10. Burattini, S. *et al. J. Am. Chem. Soc.* **132,** 12051–12058 (2010).

EVOLUTIONARY BIOLOGY

Light sense

Evidence that a larval brachiopod has ciliary photoreceptors that are directionally selective, and therefore may function as eyes, bears on an enduring puzzle about photoreceptor evolution in animals.

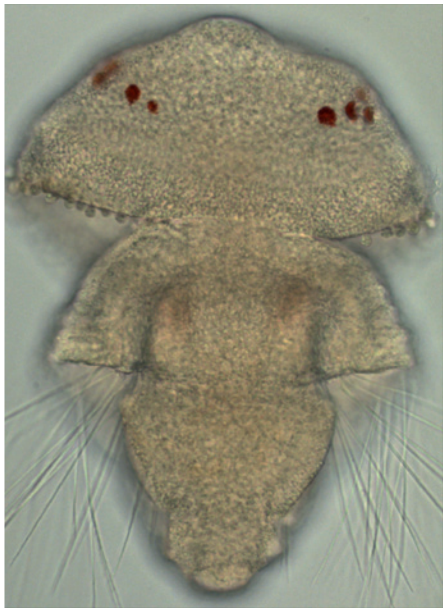
DANIEL OSORIO

Brachiopods are bivalved marine animals that superficially resemble molluscs. They are an ancient lineage, and the genus *Lingula,* which appeared 400 million years ago, still thrives. It is a pleasure of modern biology that the understanding of cellular processes built on a few model organisms is now extending to groups such as brachiopods that hitherto were scarcely known in the laboratory. Often, the richer view of evolutionary variation that emerges yields insight into how physiological mechanisms are adapted to biological needs.

An intriguing example now comes in a paper by Passamaneck *et al.* in *EvoDevo*[1], describing studies of the planktonic larva of a brachiopod named *Terebratalia* (Fig. 1). They find that the larva violates a distinction between the type of visual photoreceptor that is used by chordates (including vertebrates such as ourselves) and that used by protostomes (such as arthropods, annelids, molluscs — and brachiopods)[2,3].

Animals have two kinds of photoreceptor cells that use rhodopsin photopigments: rhabdomeric and ciliary. In each case, light activates rhodopsin, but the subsequent stages of phototransduction are quite different. One consequence is that in rhabdomeric receptors light causes cation channels to open in the photoreceptor cell membrane, thereby depolarizing the cell, whereas in ciliary receptors light hyperpolarizes the cell by closing these channels.

The common ancestor of extant bilaterally symmetrical animals had both ciliary and rhabdomeric photoreceptors, and they continue to coexist in living phyla, including chordates, molluscs and annelids[2]. Is the evolutionary persistence of two photoreceptor types indicative of their being adapted to different roles, or simply of developmental conservatism[3]? A particular puzzle is that visual photoreceptors — those present in eyes — are ciliary rods and cones in vertebrates, but rhabdomeric in protostomes. Conversely, vertebrate rhabdomeric receptors have non-visual roles such



N. FURCHHEIM, BERLIN MUS. NAT. HIST.

Figure 1 | The *Terebratalia* larva. The eyespots (red) contain photoreceptors that have cilia and ciliary rhodopsin. On the evidence of Passamaneck and colleagues' anatomical and genetic studies[1], these photoreceptors function as eyes rather than as ambient-light detectors. The larva is about 200 micrometres long.

as regulating circadian rhythms, whereas protostome ciliary receptors have, until now, been thought to be non-visual[2,3].

Brachiopods belong to the superphylum Lophotrochozoa, which includes annelids and molluscs, and so were expected to have rhabdomeric visual photoreceptors[2]. However Passamaneck and co-workers[1] find that the *Terebratalia* larva, although barely 200 micrometres long, has photoreceptors containing cilia and ciliary rhodopsin. There is no sign of a rhabdomeric rhodopsin gene.

The interest here rests on these photoreceptors acting as eyes, rather than as ambient-light detectors. Each of the larva's three to eight pairs of eyespots consists of two photoreceptor cells, one of which also contains screening pigment and the other a lens-like intracellular body.

This economical arrangement should confer directional sensitivity, which is a minimal requirement of vision[4]. Also, axons connect the photoreceptor cells to the nervous system. This direct evidence for eyes is supported by the fact that the eyespots are located where *Terebratalia*'s orthologues of two homeobox genes (*Pax6* and *Otx*), widely associated with eye development, are co-expressed. In behavioural tests carried out by Passamaneck *et al.*[1] the mature larvae did not respond to light, although it would be surprising if they are indeed blind. The authors report that gastrula embryos, an earlier stage of development at which the animals lack eyes but possess rhodopsin, move towards a light source. How this directionality might be achieved is unclear.

The existence of a ciliary visual photoreceptor in the Lophotrochozoa is not unprecedented. In addition to rhabdomeric eyes, some molluscs, including scallops, have ciliary detectors that initiate shell closure in response to shadows warning of potential predators[2,3]. Such instances of ciliary vision in a predominantly rhabdomeric lineage suggest that natural selection can vary the choice of photoreceptor according to need, and hence that each type has particular advantages.

Ciliary and rhabdomeric phototransduction mechanisms are well studied, and have fundamental differences. For example, the ciliary intracellular signalling pathway uses cyclic nucleotides, whereas the rhabdomeric receptor pathway is based on phospholipase C and calcium ions[2]. But the significance of those differences is not obvious[3]. Insects have rhabdomeric receptors and vertebrates have ciliary receptors, but both have excellent vision. Comparisons of vertebrate and invertebrate eyes usually emphasize physical limitations to light sensing, and imply that the performance of their phototransduction mechanisms is similar; for example, that at low light intensities photoreceptors are near-optimal photon counters.

In this context, the work by Passamaneck *et al.*[1] complements a review by Fain *et al.*[3] of differences in photoreceptor performance. The authors[3] argue that, in bright light, ciliary receptors are superior to rhabdomeric receptors because they consume less energy and suffer less from variation in response time-course — which reduces signal reliability. Also, a higher photopigment density in ciliary receptors enhances their sensitivity. Rhabdomeric receptors, however, function over the huge intensity range from starlight to bright sun, whereas the ciliary mechanism has to trade off response speed against the rate of spontaneous photopigment activation in the absence of light[5]. This spontaneous activation, or dark noise, is equivalent to a constant veiling light, which, in cone photoreceptors, overwhelms vision at low intensities but is insignificant in daylight.

To overcome this problem, vertebrates have a duplex retina of rods and cones. Low spontaneous activity allows rods to signal detection of a single photon, but they suffer from slow responses and take many minutes to recover from an effect known as bleaching, which leaves them blind in daylight. Vertebrate cones have fast responses, but dark noise makes them useless at night. Fain *et al.*[3] suggest that the ancestral chordate had a cone-like ciliary photoreceptor and was active in bright light. The acquisition of separate rods and cones, perhaps following a genome duplication, then allowed early vertebrates to see over as wide a range of light levels as the early protostomes that used the rhabdomeric system for vision.

Taking the work of Passamaneck *et al.*[1] further, one line of study will be to find out whether brachiopod photoreceptors are rod- or cone-like, or indeed have novel properties. Also, little is known about how the planktonic *Terebratalia* larvae live, beyond the facts that they cannot feed and have to find a site to settle. Interestingly, the sessile adults occupy a great range of depths, whereas the larvae's possession of ciliary receptors perhaps implies that they inhabit shallow, sunlit waters. ■

Daniel Osorio *is in the School of Life Sciences, University of Sussex, Brighton BN1 9QG, UK.*
e-mail: d.osorio@sussex.ac.uk

1. Passamaneck, Y. J., Furchheim, N., Hejnol, A., Martindale, M. Q. & Lüter, C. *EvoDevo* **2,** 6 (2011).
2. Arendt, D. *Int. J. Dev. Biol.* **47,** 563–571 (2003).
3. Fain, G., Hardie, R. & Laughlin, S. B. *Curr. Biol.* **20,** R114–R124 (2010).
4. Nilsson, D.-E. *Phil. Trans. R. Soc. Lond. B* **364,** 2833–2847 (2009).
5. Osorio, D. & Nilsson, D.-E. *Curr. Biol.* **14,** R1051–R1053 (2004).

ATOMIC PHYSICS

A route to quantum magnetism

The trend towards using ultracold atoms as simulators of condensed-matter and many-body phenomena is gaining momentum. These systems can now be used to simulate quantum magnetism. SEE ARTICLE P.307

IAN B. SPIELMAN

At nanokelvin temperatures — 100 billion times closer to absolute zero than room temperature — ultracold atoms are the coldest stuff in the known Universe. At such low temperatures, these quantum gases, each consisting of hundreds to hundreds of millions of atoms, are quintessential quantum many-body systems. Unlike their more conventional solid-state cousins, nearly every property of quantum gases can be controlled[1] in the laboratory with unprecedented flexibility and exquisite precision*.

On page 307 of this issue, Simon *et al.*[2] create one-dimensional chains of ultracold rubidium-87 atoms that behave as if strong magnetic interactions were present[3]. Their novel technique encodes the magnetic properties into the configuration, or geometry, of the atoms in each chain instead of the atoms' quantum mechanical spins as is usually the case in magnetic systems. This increases, by more than a factor of ten, the temperature at which magnetic ordering takes place (when the spatial pattern of spins becomes ordered rather than random), finally allowing quantum magnetism to be studied using ultracold atoms.

Understanding the properties of many-particle interacting systems is a crucial component of modern science. How do interactions between relatively simple components give rise to complex behaviour? How does consciousness emerge from a collection of neurons? How is the large-scale Universe organized? Why do electrons superconduct in some materials and give rise to magnetism in others? Each of these questions involves the emergence of macroscopic properties that are qualitatively distinct from the properties of the individual components.

In physics, we are interested in understanding and classifying the generic ways in which such macroscopic phenomena emerge, linking the microscopic properties of the components (including their environment) to the macroscale. This is usually the domain of condensed-matter physics, in which innumerable interacting electrons in materials give rise to complex phenomena such as superconductivity, the quantum Hall effect and magnetic ordering. Ultracold atoms provide a unique system in which to study many-body physics because the components and their environment are simple and can be nearly perfectly controlled. Indeed, many-body phenomena studied using ultracold atoms have been subjected to unprecedented quantitative comparison with many-body theories. Examples of this

*This article and the paper under discussion[2] were published online on 13 April 2011.

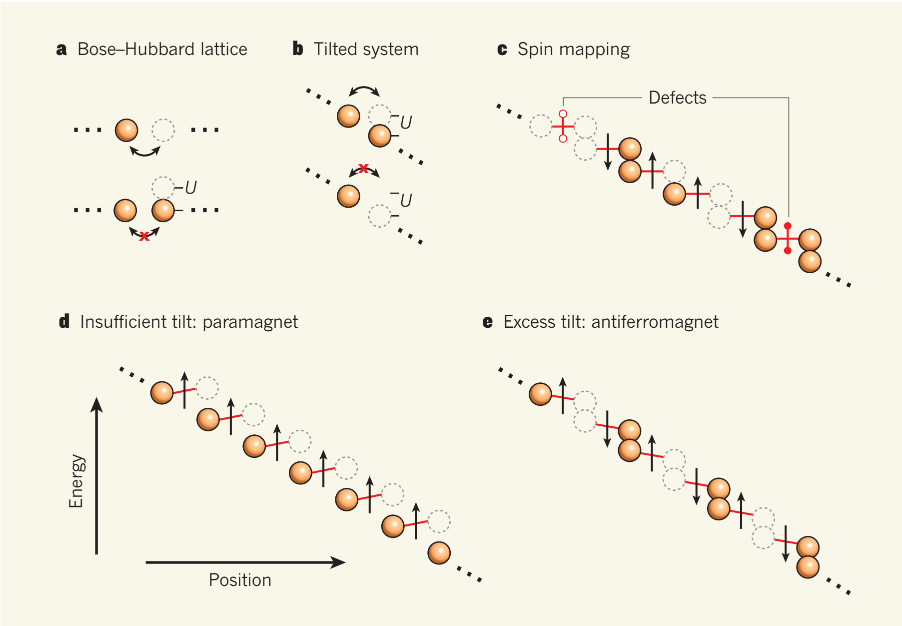


Figure 1 | The principle of Simon and colleagues' effective spin model[2]. **a**, Two examples of a Bose–Hubbard lattice: top, energetically allowed hopping from an occupied atomic site to a neighbouring empty site; bottom, suppression of hopping between neighbouring occupied sites owing to the on-site repulsive energy *U* between two atoms. **b**, Two examples of a tilted lattice. The tilt is sufficient to overcome on-site repulsion, such that the situation in **a** is reversed: top, energetically allowed hopping between neighbouring occupied sites; bottom, suppression of hopping between an occupied site and a neighbouring vacant site. **c**, Mapping of effective spin states (arrows) into different spatial configurations. Defects denote configurations that are not associated with spins, showing that the mapping is approximate. **d**, Insufficient tilt. A system with an average of one atom per site will remain in this configuration, in which the spins align parallel to one another (a paramagnet). **e**, Excess tilt. A system with an average of one atom per site will minimize its energy by creating an alternating array of empty sites and doubly occupied sites, corresponding to an arrangement in which the spins are staggered (an antiferromagnet).

are the phase transition from superfluid to Mott insulator[4], and a continuous transition known as the Bose–Einstein condensate–Bardeen–Cooper–Schrieffer (BEC–BCS) crossover[5,6].

Magnetism is a topic of fundamental importance in condensed-matter physics, having an impact on technology (such as magnetic storage media and some visions of 'spintronic' devices[7,8]) as well as on concepts at the forefront of physics (such as spin liquids[9]). Magnetic systems cannot yet be directly studied using quantum gases because their magnetic-ordering temperatures are tens or hundreds of picokelvins (1 picokelvin is 10^{-12} kelvins). So far, this is too cold even for ultracold atoms.

To bypass this limitation, Simon *et al.*[2] constructed an atomic system in which the spatial configuration of the atoms encodes the 'spin' usually associated with magnetism. They first created a rubidium-87 BEC, in which all the atoms are essentially in the same quantum state, and compressed it into a single two-dimensional (2D) plane. Next, using two independent standing waves of light, they confined the atoms into the sites of a square lattice with one atom per site. One of the two standing waves was made more intense than the other, so that the atoms could 'hop' between sites only in one-dimensional (1D) chains; such hopping is suppressed when neighbouring lattice sites are already occupied, because of the repulsive energy involved in locating two atoms on a single site (Fig. 1a).

To create a magnetic analogue, Simon *et al.*[2] superimposed a magnetic-field gradient on the optical lattice potential, creating an energy difference, or 'tilt', between neighbouring lattice sites. When this energy difference is large enough, it can exceed the repulsive interaction energy, allowing atoms to move from site to site. Thus, a sufficiently large tilt reverses the earlier picture: when a neighbouring site is occupied hopping is allowed (Fig. 1b), but when it is empty hopping is blocked (to conserve energy). When the tilt increases from zero, 1D chains of atoms undergo a transition from one atom per site (Fig. 1d; hopping blocked, analogous to a paramagnet) to a staggered, alternating array of empty sites and doubly occupied sites (Fig. 1e; analogous to an antiferromagnet). Here, the effective magnetism relies on repulsion between atoms, rather than the much smaller exchange interactions between the physical spin states, which normally lead to magnetism in materials. Because every aspect of this system is well understood, the authors were able to directly compare their results with the 1D quantum Ising magnetic model they sought to realize — and found spectacular agreement.

Like the Brownian motion of pollen grains suspended in water, physical systems are usually coupled to a source of random thermal motion: an environment. The successful implementation of this technique relies heavily on a second feature of ultracold atom systems: they are not well connected to such an environment. When the magnetic-field gradient tilts the lattice potential, essentially creating a hill for the atoms, one might expect the atoms to reduce their energy by all moving down the hill. Because the atoms are not linked to an environment, there is no mechanism for them to begin this downhill motion and nowhere for the excess energy to go, so instead the system remains within the subset of spatial configurations that are analogous to spins.

Simon and colleagues[2] demonstrate a powerful experimental technique for creating analogues to spin systems. However, the 1D quantum Ising model is well understood from modern theoretical techniques[10], and the small systems studied here can be numerically exactly solved by brute force. It is the next step that is profound: creating spin systems that are currently insoluble by known analytical or numerical techniques. For example, the extensions of this work discussed by Sachdev *et al.*[3] would allow the simulation of 2D 'frustrated' spin systems, for some of which an exotic, ground (minimal-energy) state of matter known as a quantum spin liquid is expected[9]. As the complexity and control of ultracold atoms have continued to increase, it has become evident that they are not simply emulators of condensed-matter systems, but fascinating quantum many-body systems in their own right. As such, they may well exhibit behaviours that do not exist in any other physical system. Extensions of Simon and colleagues' technique[2] should allow the construction of nearly arbitrary spin systems, which will be a crucial leap in that direction. ■

Ian B. Spielman *is at the Joint Quantum Institute, National Institute of Standards and Technology and the University of Maryland, Gaithersburg, Maryland 20899–8424, USA.*
e-mail: ian.spielman@nist.gov

1. Bloch, I., Dalibard, J. & Zwerger, W. *Rev. Mod. Phys.* **80,** 885–964 (2008).
2. Simon, J. *et al. Nature* **472,** 307–312 (2011).
3. Sachdev, S., Sengupta, K. & Girvin, S. M. *Phys. Rev. B* **66,** 075128 (2002).
4. Greiner, M. *et al. Nature* **415,** 39–44 (2002).
5. Regal, C. A., Greiner, M. & Jin, D. S. *Phys. Rev. Lett.* **92,** 040403 (2004).
6. Zwierlein, M. W. *et al. Phys. Rev. Lett.* **92,** 120403 (2004).
7. Žutić, I., Fabian, J. & Das Sarma, S. *Rev. Mod. Phys.* **76,** 323–410 (2004).
8. Ralph, D. C. & Stiles, M. D. *J. Magn. Magn. Mater.* **320,** 1190–1216 (2008).
9. Balents, L. *Nature* **464,** 199–208 (2010).
10. Ovchinnikov, A. A., Dmitriev, D. V., Krivnov, V. Ya. & Cheranovskii, V. O. *Phys. Rev. B* **68,** 214406 (2003).

CANCER

Macrophages limit chemotherapy

A major hurdle to successful cancer treatment is tumour resistance to chemotherapy. White blood cells called macrophages often infiltrate tumours in large numbers, and now appear to promote tumour chemoresistance.

MICHELE DE PALMA & CLAIRE E. LEWIS

Malignant tumours are rich in immune cells, including macrophages and neutrophils[1]. It was initially thought that the body recruits such cells to tumours in order to defend itself against the developing cancer. However, there is increasing evidence for the opposite being true. For instance, the tumour microenvironment seems to suppress the normal anticancer functions of both macrophages and another type of immune cell, the cytotoxic T lymphocyte[2,3]. Even worse, macrophages promote, rather than restrict, tumour progression by stimulating tumour vascularization, invasion and metastasis[3]. Now, in a fascinating paper published in *Cancer Discovery*, DeNardo *et al.*[4] show that these cells can also enhance tumour resistance to chemotherapy.

Macrophages and neutrophils are normally at the front line of the body's defence: they engulf pathogens and dead or dying cells, and repair wounds. These white blood cells (leukocytes) also function in tumours, along with cells that determine the specificity of the immune response against foreign agents — T and B lymphocytes. Previous work by the same team[4] showed that subpopulations of leukocytes, including macrophages, communicate with one another in tumours to promote malignancy. In mouse models of cancer, for example, the CD4+ helper T lymphocytes and B lymphocytes release interleukin-4 (an immune mediator) and antibodies, respectively, to stimulate the tumour-promoting functions of macrophages[5,6].

The first hint that the function of these tumour-associated macrophages (TAMs) and lymphocytes might be linked in human tumours comes from the present study[4]. DeNardo *et al.* identified a distinct immune-cell 'signature' in surgically removed breast tumours that predicted the survival of patients: women with a combination of high TAMs, high CD4+ helper T cells and low cytotoxic T cells in their tumours were at high risk of developing secondary tumours and, therefore, of succumbing to the disease. Earlier studies had established a link between poor prognosis and either high TAM numbers[3,7] or low T-cell numbers[8] in tumours. What DeNardo *et al.* show for the first time is the pronounced prognostic value of the inverse correlation between numbers of TAMs and cytotoxic T cells.

The researchers[4] also show that breast tumours with high TAM numbers and low numbers of cytotoxic T cells respond relatively poorly to chemotherapy given before surgery. This may be because chemotherapy stimulates tumour cells to release a protein called CSF1. Indeed, in mouse mammary tumours, chemotherapy increased tumour-cell expression of CSF1, which then recruited large numbers of macrophages expressing the CSF1 receptor to the tumour. This effect may depend on the chemotherapeutic agent used, the type of tumour, and even the immune status of the host. For example, in a previous study[9], a different form of chemotherapy failed to increase TAM numbers in human breast tumours grown in mice lacking T cells.

Pharmacological blockade of macrophage recruitment markedly improved the ability of the chemotherapeutic agent paclitaxel to slow the growth of both primary and metastatic tumours (Fig. 1). DeNardo and colleagues suggest that the ability of TAMs to limit tumour responses to chemotherapy is, at least in part, due to their suppression of the antitumour functions of cytotoxic T cells. Their paper therefore attests to the functional complexity of the immune-cell repertoire in tumours and shows that crosstalk between TAMs and cytotoxic T cells affects tumour responses to chemotherapy. Nonetheless, more work is required to better dissect this phenomenon and its therapeutic relevance.

TAMs themselves represent a heterogeneous assortment of functionally distinct cells[3]. Recent studies[3,10] in mice have identified and characterized various TAM subsets by their gene-expression signatures and specific location in distinct tumour regions. Furthermore, genetic tools have helped to elucidate specific functions of these TAM subsets[3,10,11].

Intriguingly, DeNardo *et al.* report that blocking CSF1 receptors selectively inhibits TAM infiltration into tumour regions distant from the blood vessels, leaving perivascular TAMs apparently unaffected. Because this partial depletion of TAMs improved the efficacy of chemotherapy, it is conceivable that the TAMs present in poorly vascularized areas specifically protect tumours from the effects of chemotherapy. Further studies are warranted to characterize this subset of TAMs and to see whether their exposure to distinct signals in

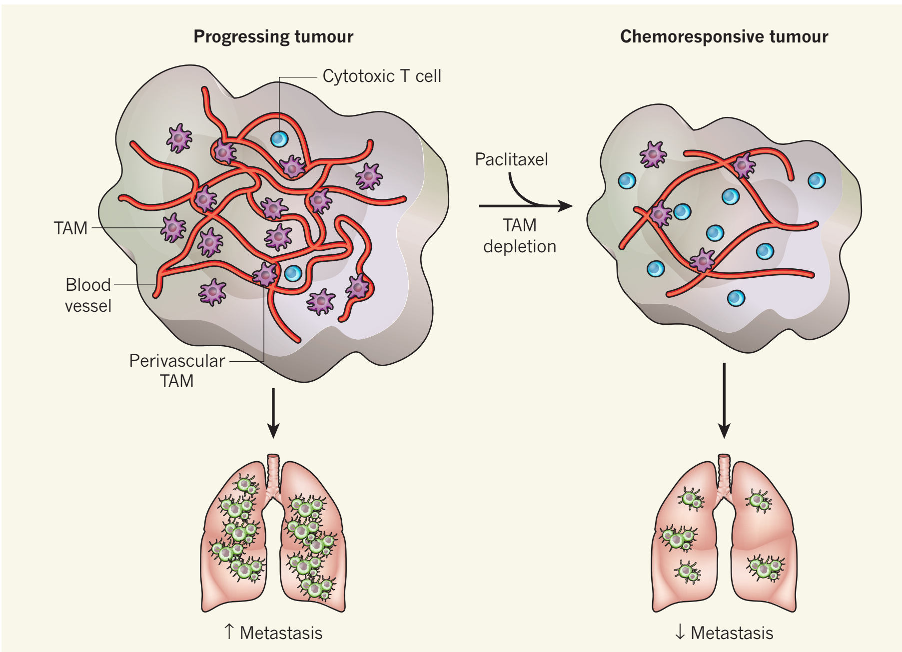


Figure 1 | Macrophages promote tumour chemoresistance. Progressing mammary tumours contain a variety of leukocytes, including tumour-associated macrophages (TAMs) and cytotoxic T cells. DeNardo *et al.*[4] report that TAM depletion in such tumours increased the antitumour efficacy of the chemotherapeutic agent paclitaxel. Consequently, paclitaxel treatment resulted in a greater reduction in tumour size in TAM-depleted tumours, which also contained fewer blood vessels, higher numbers of cytotoxic T cells and more signs of tumour destruction. It also reduces the extent of metastasis in organs such as the lung.

50 Years Ago

In *Nature* of October 12, 1957 … we were glad and proud to congratulate the scientists and technologists of the U.S.S.R. on the successful launching of the first artificial Earth satellite (*Sputnik I*). Less than a month afterwards … we again had occasion to congratulate the U.S.S.R., when the second satellite (*Sputnik II*) was launched, this time containing a dog ... Now again, on April 12, a still further point in scientific advancement was reached, for on that day scientists and technologists in the U.S.S.R. reaped a rich reward for concentrated and ingenious research by successfully launching another space-vehicle (*Vostok*) into space. This satellite carried, among other things, a man (Major Yuri Alekseyevich Gagarin), travelled in orbit around the Earth in 89 min. 6 sec., and a little later was brought back safely to Earth, with Major Gagarin unharmed — all in about 108 min.
From *Nature* 22 April 1961

100 Years Ago

The first non-stop flight from London to Paris was made, on April 12, on a Blériot monoplane by M. Pierre Prier in 3h. 56m. M. Prier, who is the chief instructor of the Blériot School at Hendon, left that ground at 1.37 p.m., taking a course for Dover via Hampstead, Highgate, Greenwich, Chatham, and Canterbury … [H]e reached Dover, at 2.50 p.m. Thirty minutes later he was over Boulogne, and steering a straight course over Abbeville and Beauvais for Paris, where he arrived at 5.33 p.m., making a perfect landing in front of the Blériot sheds at the Issy-les-Moulineaux aviation ground … M. Prier … suffered no inconvenience throughout the journey except slight inflammation of the eyes, due to his neglecting to equip himself with goggles.
From *Nature* 20 April 1911

these tumour sites stimulates their chemoprotective functions. It will then be possible to investigate whether a similar TAM subset exists in human breast tumours — and in other types of tumour.

The authors also found that the combination of TAM depletion and chemotherapy reduced tumour-vessel density by 50% and led to greater tumour destruction. TAMs are known to stimulate excessive tumour vascularization[3,7,10,12], and their persistent expression of pro-angiogenic factors such as VEGF induces the formation of abnormal, hypo-perfused blood vessels, which limit the delivery of chemotherapy to tumours[10,11]. It is therefore conceivable that TAM depletion at sites away from blood vessels[4] may have skewed the perivascular TAMs from a pro-angiogenic[3,10] to an angiostatic function[12], thereby 'normalizing' the remaining vessels to acquire the structure and function of vessels in healthy tissues. This would temporarily increase blood flow to the tumour, enhancing the delivery and so the efficacy of the chemotherapeutic agent.

The findings of DeNardo and colleagues[4] add weight to an emerging compelling case for deciphering the complexity of leukocyte infiltrates in breast cancer. Their findings suggest that this may provide clinically relevant indications of the likely response to chemotherapy and thus patient survival. Future studies should show the relevance of the chemoprotective subset of TAMs that these authors identify for human cancer, and may even highlight molecular targets for therapies that restrain the activity of these cells or even reactivate their antitumour functions. Such reprogramming of the immune microenvironment in tumours is a promising strategy for improving the efficacy of standard anticancer treatments. ■

Michele De Palma *is at the Angiogenesis and Tumour Targeting Research Unit, and HSR-TIGET, San Raffaele Scientific Institute, 20132-Milan, Italy.* **Claire E. Lewis** *is at the Academic Unit of Tumour Inflammation & Tumour Targeting, Department of Oncology, University of Sheffield Medical School, Sheffield S10 2RX, UK.*
e-mails: depalma.michele@hsr.it; claire.lewis@sheffield.ac.uk

1. de Visser, K. E., Eichten, A., Coussens, L. M. *Nature Rev. Cancer* **6,** 24–37 (2006).
2. Mellor, A. L. & Munn, D. H. *Nature Rev. Immunol.* **8,** 74–80 (2008).
3. Lewis, C. E. & Pollard, J. W. *Cancer Res.* **66,** 605–612 (2006).
4. DeNardo, D. G. *et al. Cancer Discov.* doi:10.1158/2159-8274.CD-10-0028 (2011).
5. DeNardo, D. G., Andreu, P. & Coussens, L. M. *Cancer Metastasis Rev.* **29,** 309–316 (2010).
6. Andreu, P. *et al. Cancer Cell* **17,** 121–134 (2010).
7. Bingle, L., Brown, N. J. & Lewis, C. E. *J. Pathol.* **196,** 254–265 (2002).
8. Galon, J. *et al. Science* **313,** 1960–1964 (2006).
9. Paulus, P., Stanley, E. R., Schäfer, R., Abraham, D. & Aharinejad, S. *Cancer Res.* **66,** 4349–4356 (2006).
10. Pucci, F. *et al. Blood* **114,** 901–914 (2009).
11. Stockmann, C. *et al. Nature* **456,** 814–818 (2008).
12. Rolny, C. *et al. Cancer Cell* **19,** 31–44 (2011).

ELECTRONICS

Power surfing on waves

Wavy strips of piezoelectric materials on stretchable substrates can both withstand larger applied mechanical strain without cracking and harvest energy more efficiently than their flat counterparts.

MIN HYUNG LEE & ALI JAVEY

Bendable and stretchable electronics such as roll-up keyboards and displays have garnered tremendous attention because of their portability and versatility. However, a major limitation of all portable devices available today is their short battery life, which often limits their functional use to just a few hours after a full charge. If an electronic device could be self-powered — for instance, by harvesting energy from the mechanical deformation it undergoes during typing, rolling and unrolling — then the power problem would be greatly reduced. Writing in *Nano Letters*, Qi *et al.*[1] describe a technique to produce thin ribbons of wavy piezoelectric materials on rubber substrates that paves the way to the realization of highly stretchable, energy-harvesting devices.

In piezoelectric materials, applied mechanical strain results in the accumulation of electric charge and the generation of electricity. This phenomenon, known technically as the piezoelectric effect, was first observed in 1880 by the French physicists and brothers Pierre and Jacques Curie. Since then, the effect has been used in a wide range of applications, including sensors and energy-harvesting devices. A few popular examples of inorganic piezoelectric materials include zinc oxide, aluminium nitride and lead zirconate titanate (PZT).

In their bulk form, these inorganic semiconductors are mechanically rigid. But once miniaturized to strips or wires of nanometre-scale thickness, they become mechanically flexible. Stretchability, however, is more challenging. In the past, researchers have demonstrated[2] that buckled silicon ribbons with periodic, wavy (sinusoidal) shapes can accommodate extreme increases in maximum

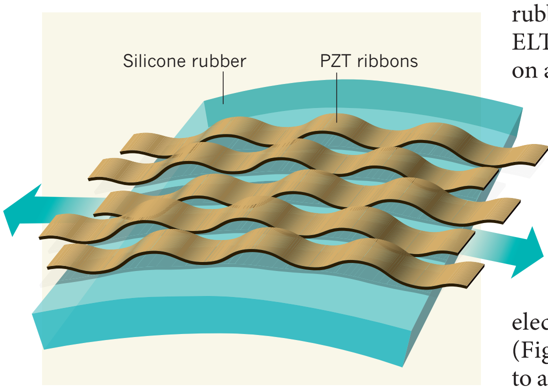


Figure 1 | The design of Qi and colleagues' energy-harvesting devices[1]. The stretchable devices consist of wavy ribbons made of piezoelectric material (lead zirconate titanate; PZT) on silicone rubber. They can withstand greater applied mechanical strain without cracking than can equivalent materials made into flat ribbons.

strain without fracturing. Qi *et al.*[1] have now extended this concept for the fabrication of stretchable, energy-harvesting devices using the technique of epitaxial layer transfer (ELT).

ELT involves the selective etch, lift-off and transfer of a single-crystalline thin film from a primary substrate, on which the film has been grown, to a secondary, receiver substrate. The technique allows the placement of a user-defined semiconductor 'X' on a substrate 'Y' (termed X-on-Y, or XoY) and has proved useful in a broad spectrum of electronic and photonic applications. ELT was used[3,4] as early as 1975, when micrometre-scale gallium arsenide/aluminium gallium arsenide (GaAs/AlGaAs) films were transferred from bulk GaAs wafers to glass substrates to enable the production of GaAs night-vision goggles and solar cells. In the late 1980s, a similar process was used[5] to fabricate thin-film GaAs diode lasers and other optoelectronic devices on glass substrates. And in the mid-to-late 2000s, researchers demonstrated[2,6] the ability to extend the ELT technique to substrates that are mechanically flexible and stretchable. This then led to the development of electronic devices that can conform to cover irregular and curvy surfaces.

Recently, the utility of ELT was further expanded[7] to allow the transfer of ultrathin films of type III–V semiconductors — down to a few nanometres in thickness — to silicon/silicon oxide (Si/SiO_2) substrates. This advance[7] offers the advantage of combining III–V semiconductors, which have high electron mobility, with well-established silicon technology for the exploration of energy-efficient electronics. Qi and colleagues' work[1] demonstrates yet another application of the ELT process: the production of buckled PZT strips on a rubber substrate for use as piezoelectric energy-harvesting devices.

More specifically, their approach is as follows. First, they pre-stretched a slab of silicone rubber. Second, they transferred, by means of ELT, PZT ribbons that had been patterned on a magnesium oxide substrate to the pre-stretched silicone substrate. After releasing the rubber substrate, the initially flat ribbons buckled, forming wavy geometries of amplitude and wavelength that depended on, among other factors, the thickness of the ribbons and their interaction with the rubber.

Next, the authors measured the piezoelectric response of the wavy PZT ribbons (Fig. 1), as well as that of flat PZT ribbons, to applied mechanical strain. In comparison to flat PZT ribbons, which exhibited failure under less than 1% strain, wavy ribbons withstood up to about 8% strain without cracking. Interestingly, wavy PZT ribbons also displayed an enhancement of up to 70% in piezoelectric response compared with flat ribbons, which the authors attribute to the large location-dependent strain gradient of the wavy structures.

Qi and colleagues' stretchable, piezoelectric energy-harvesting materials are an important advance in the rapidly growing field of conformal electronics and sensors. But as they themselves note[1], as far as the actual manufacturing of devices is concerned, a number of engineering challenges remain. Specifically, low-cost generation of wavy piezoelectric strips on large areas of substrate — ranging from cubic centimetres to metres of surface coverage — has yet to be demonstrated, and requires further research. In addition, more work is needed to determine the maximum density of electrical energy that can be generated by the devices, which will effectively establish their potential niche applications. ■

Min Hyung Lee *and* **Ali Javey** *are in the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720, USA.*
e-mail: ajavey@berkeley.edu

1. Qi, Y. *et al. Nano Lett.* **11,** 1331–1336 (2011).
2. Sun, Y., Choi, W. M., Jiang, H., Huang, Y. Y. & Rogers, J. A. *Nature Nanotechnol.* **1,** 201–207 (2006).
3. Antypas, G. A. & Edgecumbe, J. *Appl. Phys. Lett.* **26,** 371–372 (1975).
4. Konagai, M., Sugimoto, M. & Takahashi, K. *J. Cryst. Growth* **45,** 277–280 (1978).
5. Yablonovitch, E., Gmitter, T., Harbison, J. P. & Bhat, R. *Appl. Phys. Lett.* **51,** 2222–2224 (1987).
6. Kim, D.-H. *et al. Science* **320,** 507–511 (2008).
7. Ko, H. *et al. Nature* **468,** 286–289 (2010).

IMMUNOLOGY

TRIM5 does double duty

TRIM5 proteins limit retroviral infection by targeting the viral coat. It now seems that these proteins can also serve as pattern-recognition receptors, which initiate cellular innate immune responses. SEE LETTER P.361

CHRISTOPHER AIKEN & SEBASTIAN JOYCE

Mammalian proteins use various mechanisms to restrict retroviral infection[1]. One such protein, TRIM5, blocks HIV-1 infection in some primate species: when the virus enters a cell, TRIM5 engages with the viral coat, or capsid, inducing premature uncoating[2]. On page 361 of this issue, Pertel and co-workers[3] report that this protein has another function: TRIM5 is involved in activating cellular innate immune responses.

Cells sense infection with viruses and bacteria by detecting pathogen-specific molecules, including double-stranded RNA, lipids and carbohydrates. These molecules often form repeating structures — referred to as pathogen-associated molecular patterns — that bind to pattern-recognition receptors and activate signalling pathways that result in inflammation and cellular resistance to infection.

Pertel *et al.* show that increased expression of TRIM5 promotes the expression of specific genes by activating immune signalling pathways that are mediated by the transcription factors AP-1 and NF-κB. Accordingly, the removal of TRIM5 from cells not only reduces the expression of specific genes involved in innate cellular responses, but also allows HIV-1 to infect cells exposed to LPS. (LPS is a bacterial membrane component that binds the pattern-recognition receptor TLR4 and prevents HIV-1 infection by inducing a cellular antiviral state.) TRIM5's contribution to LPS-induced resistance to HIV-1 is probably independent of viral recognition: the cells Pertel and colleagues studied contain human TRIM5α, which does not bind the HIV-1 capsid efficiently. Nonetheless, their data indicate that TRIM5 contributes to innate immune responses triggered by a specific pattern-recognition receptor.

In addition to binding to and restricting retroviral capsids, TRIM5 is an E3 ligase enzyme that catalyses the attachment of the small modifier protein ubiquitin to itself and, potentially, to other proteins[4]. Whether the ubiquitin-ligase activity of TRIM5 plays a part in its retrovirus-restriction activity had remained unclear.

The attachment of multiple ubiquitin molecules to a target protein often leads to

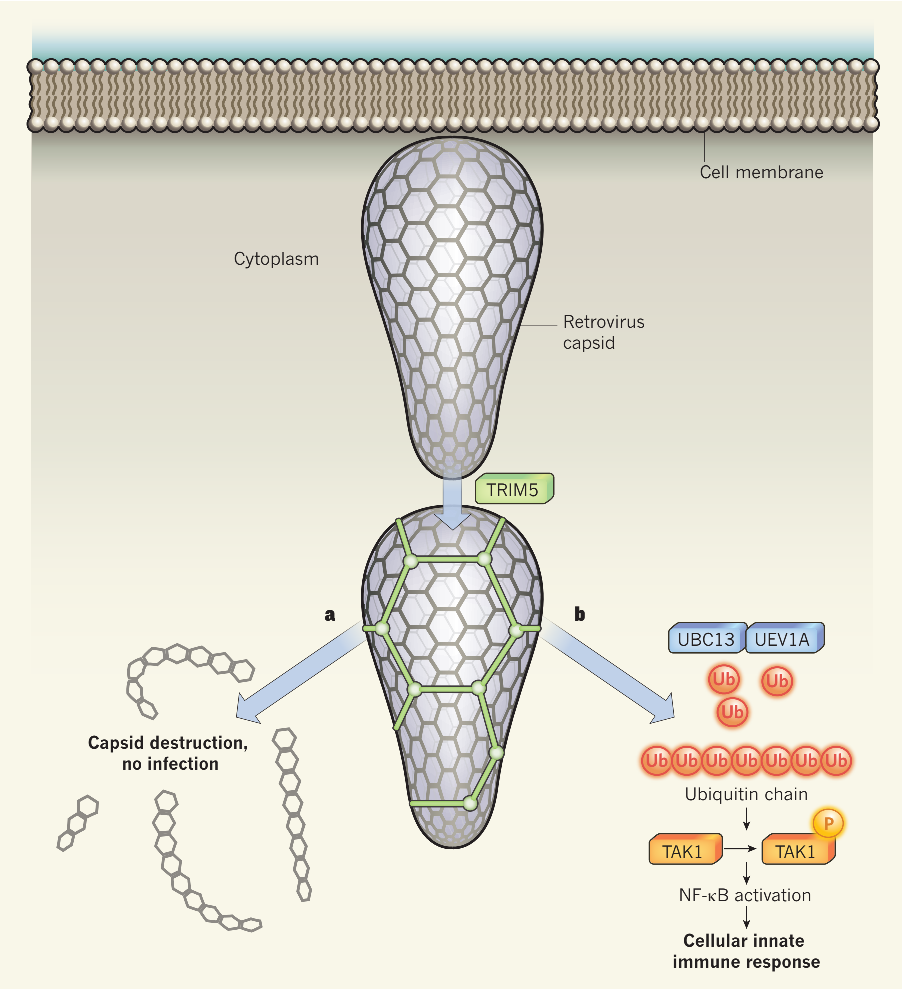


Figure 1 | TRIM5 strikes twice. When a retrovirus enters the host cell cytoplasm, TRIM5 proteins bind to the viral capsid, causing structural alterations that lead to capsid fragmentation (**a**). Pertel *et al.*[3] show that TRIM5 binding to the viral capsid also activates the protein's ubiquitin ligase activity, which, together with the UBC13–UEV1A enzyme complex, results in the synthesis of ubiquitin (Ub) chains. The chains stimulate TAK1 phosphorylation and the expression of NF-κB-dependent genes (**b**). This may induce an antiviral state, further protecting the host from infection.

its destruction. But ubiquitin also has other cellular roles. For instance, unattached ubiquitin chains, which form through the binding of one ubiquitin molecule to the lysine 63 (K63) amino-acid residue of another ubiquitin, can directly activate protein kinase enzymes[5] such as TAK1, which mediates innate immune signalling by activating AP-1 and NF-κB. In a survey of candidate host proteins through which TRIM5 might signal, Pertel *et al.*[3] found that this protein associates with TAK1 and with TAB2 and TAB3 — adaptor proteins that facilitate TAK1 activation. The ubiquitin-conjugating enzyme complex UBC13–UEV1A also associates with TRIM5 (Fig. 1).

To analyse the catalytic activity of TRIM5 *in vitro*, the authors produced TRIM5–CypA, a fusion protein, from owl monkey cells. This is a noteworthy technical achievement, because TRIM5 proteins are notorious for their insolubility and so are difficult to manipulate. Using this construct, Pertel *et al.* demonstrated that TRIM5 catalyses the formation of K63-linked ubiquitin chains *in vitro*, activating TAK1. TRIM5 extracted from human cells also made K63-linked ubiquitin chains, suggesting that this activity is common to TRIM5 proteins of different species. In this respect, TRIM5 proteins thus resemble TRAF6, another ubiquitin ligase that catalyses the formation of K63-linked ubiquitin chains and activates NF-κB signalling[6].

What is the link between TRIM5 binding to a retroviral capsid and its ubiquitin-ligase activity, which induces cellular innate immunity? TRIM5α assembles into a large hexameric lattice on the viral capsid surface[7], but whether capsid engagement influences its catalytic activity was unknown. In an experimentally well-controlled biochemical study, Pertel *et al.* show that TRIM5–CypA binding to artificial capsid-like structures markedly activates the synthesis of K63-linked ubiquitin chains *in vitro*. So it seems that sensing of the viral capsid induces TRIM5's ubiquitin ligase activity, which, in turn, promotes innate immunity.

The researchers did not directly investigate whether this is also the case in cells. They show, however, that exposure of the cells to a restriction-sensitive retrovirus results in the activation of several genes involved in innate immunity. For at least one of the genes, TRIM5 depletion prevented its activation. Thus, engagement of the viral capsid by TRIM5 correlates with innate immune signalling, leading the authors to conclude that this protein is a bona fide pattern-recognition receptor for retroviral capsids.

If viral engagement by the restrictive TRIM5 proteins is enough to prevent infection, what is the point of an innate antiviral response on top of this? Perhaps it is that, because TRIM5 restriction can be overcome at high viral doses, the innate response protects neighbouring cells, thereby preventing viral spread within the host. Moreover, restriction may be coupled to the signalling activity of TRIM5: Pertel *et al.*[3] report that depletion of cellular UBC13 and TAK1 enhances the retroviral infection of cells that express a restrictive TRIM5 protein. This observation is unexpected, because it has been reported that TRIM5 inflicts direct structural damage on the viral capsid[8,9] and that ubiquitination is not necessary for the restriction of viral infection[10].

It is possible that the ubiquitin ligase activity of TRIM5 is important for both restriction and cell signalling, but that the two effects are independent. Furthermore, UBC13 and TAK1 may act in other ways to enhance the restriction activity of TRIM5 — for instance, by modulating its stability, activity or intracellular localization. Whether the human TRIM5α protein, which unfortunately does not prevent HIV-1 infection, alerts the host to the presence of the virus remains to be seen.

Pertel *et al.*[3] uncover a surprising facet of TRIM5 biology and a new mechanism for cellular sensing of viruses. Their paper is sure to stimulate the interest of virologists and immunologists alike. ■

Christopher Aiken *and* **Sebastian Joyce** *are in the Department of Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA.*
e-mail: chris.aiken@vanderbilt.edu

1. Wolf, D. & Goff, S. P. *Annu. Rev. Genet.* **42,** 143–163 (2008).
2. Stremlau, M. *et al. Proc. Natl Acad. Sci. USA* **103,** 5514–5519 (2006).
3. Pertel, T. *et al. Nature* **472,** 361–365 (2011).
4. Xu, L. *et al. Exp. Cell Res.* **288,** 84–93 (2003).
5. Xia, Z.-P. *et al. Nature* **461,** 114–119 (2009).
6. Deng, L. *et al. Cell* **103,** 351–361 (2000).
7. Ganser-Pornillos, B. K. *et al. Proc. Natl Acad. Sci. USA* **108,** 534–539 (2011).
8. Langelier, C. R. *et al. J. Virol.* **82,** 11682–11694 (2008).
9. Zhao, G. *et al. PLoS Pathog.* **7,** e1002009 (2011).
10. Perez-Caballero, D., Hatziioannou, T., Zhang, F., Cowan, S. & Bieniasz, P. D. *J. Virol.* **79,** 15567–15572 (2005).

ARTICLE

doi:10.1038/nature09994

Quantum simulation of antiferromagnetic spin chains in an optical lattice

Jonathan Simon[1], Waseem S. Bakr[1], Ruichao Ma[1], M. Eric Tai[1], Philipp M. Preiss[1] & Markus Greiner[1]

Understanding exotic forms of magnetism in quantum mechanical systems is a central goal of modern condensed matter physics, with implications for systems ranging from high-temperature superconductors to spintronic devices. Simulating magnetic materials in the vicinity of a quantum phase transition is computationally intractable on classical computers, owing to the extreme complexity arising from quantum entanglement between the constituent magnetic spins. Here we use a degenerate Bose gas of rubidium atoms confined in an optical lattice to simulate a chain of interacting quantum Ising spins as they undergo a phase transition. Strong spin interactions are achieved through a site-occupation to pseudo-spin mapping. As we vary a magnetic field, quantum fluctuations drive a phase transition from a paramagnetic phase into an antiferromagnetic phase. In the paramagnetic phase, the interaction between the spins is overwhelmed by the applied field, which aligns the spins. In the antiferromagnetic phase, the interaction dominates and produces staggered magnetic ordering. Magnetic domain formation is observed through both *in situ* site-resolved imaging and noise correlation measurements. By demonstrating a route to quantum magnetism in an optical lattice, this work should facilitate further investigations of magnetic models using ultracold atoms, thereby improving our understanding of real magnetic materials.

Quantum spins arranged on a lattice and coupled to one another through magnetic interactions constitute a paradigmatic model system in condensed matter physics. Such systems produce a rich array of magnetically ordered ground states, such as paramagnets, ferromagnets and antiferromagnets. Certain geometries and interactions induce competition between these orderings in the form of frustration, resulting in spin liquids[1] and spin glasses[2], as well as phases with topological order[3]. Varying system parameters can induce quantum phase transitions between the various phases[4]. A deeper understanding of the competition and resulting transitions between magnetic phases would provide valuable insights into the properties of complex materials such as high-temperature superconductors[5], and more generally into the intricate behaviours that can emerge when many simple quantum mechanical objects interact with one another.

Studying quantum phase transitions of magnetic condensed matter systems is hindered by the complex structure and interactions present in such systems, as well as by the difficulty of controllably varying system parameters. With a few notable exceptions[6,7], these issues make it difficult to capture the physics of such systems with simple models. Accordingly, there is a growing effort under way to realize condensed matter simulators using cold atom systems[8,9], which are understood from first principles. The exquisite control afforded by cold atom experiments permits tuning of such systems through quantum phase transitions[9,10], thereby enabling investigations of criticality[11,12] and scaling[13]. Time-resolved local readout[14–16] and manipulation[17] provide direct access to local dynamics and correlations. With this powerful toolbox available, considerable attention has turned to understanding magnetic phase transitions using cold atom quantum simulations.

Initial experimental efforts to simulate quantum magnetism have focused on bulk itinerant systems of ultracold fermions[18] and small, highly connected spin-networks simulated with ion chains[19,20]. Polar molecules[21] and Rydberg atoms[22] have been the subject of preliminary investigations, both experimentally and theoretically[23–25], as alternatives to ground-state atoms with stronger, longer-range interactions. There has also been initial success in detecting ordered states, which are artificially prepared through patterned loading[17,26,27] and double-well experiments[28].

In this work, we simulate a one-dimensional chain of interacting Ising spins by using a Mott insulator[10,29,30] of spinless bosons in a tilted optical lattice[31]. Each Ising spin maps onto the motional degree of freedom of a single atom in the lattice. The more commonly considered approach instead represents the magnetic spins by two internal states of the atoms, with spin–spin interactions arising from super-exchange[32]. Super-exchange interactions in cold atoms are quite weak, though they have been successfully observed in double-well systems[33]. The approach presented here has the benefit of a dynamical timescale set by the tunnelling rate t, rather than by the super-exchange interaction t^2/U, where U is the on-site interaction energy. Combining the faster dynamics with the high spatial resolution afforded by a quantum gas microscope[14], we directly observe transitions between paramagnetic and antiferromagnetic phases as spin–spin interactions compete with magnetic fields.

One of the primary concerns in studying magnetism with ultracold atoms is the apparent difficulty of reaching the requisite temperatures[34–37]. Ultracold gases, however, are effectively isolated from their environment, and as such it is entropy and not temperature which is in principle conserved as system parameters are tuned. Spin-polarized Mott insulators have been demonstrated with defect densities approaching the per cent level[15,16], corresponding to configurational entropy far below the spin entropy required for magnetic ordering (see Supplementary Information). This allows us to use such a Mott insulator to initialize a magnetic system with low spin entropy. We engineer a magnetic Hamiltonian whose paramagnetic ground state overlaps well with the initial Mott state, and subsequently tune it through a quantum phase transition[38] to produce an antiferromagnet. The difficulty of cooling lattice spins is thus replaced with that of performing sufficiently slow ramps to minimize diabatic crossings of many-body energy gaps.

[1]Department of Physics, Harvard University, 17 Oxford Street, Cambridge, Massachusetts 02138, USA.

Ising interactions in a tilted optical lattice

The quantum Ising model is a textbook model of magnetism, and an Ising chain is one of the simplest many-body systems to exhibit a quantum phase transition. The Hamiltonian describing a one-dimensional antiferromagnetic Ising chain in the presence of an applied magnetic field is given by:

$$H = J\sum_i S_z^i S_z^{i+1} - h_z^i S_z^i - h_x^i S_x^i$$

Here $S_z^i(S_x^i)$ is the $z(x)$ spin-projection operator at site i, and $h_z^i(h_x^i)$ is the $z(x)$-component of the magnetic field applied to site i. The zero-temperature phase diagram of the model[39,40], shown in Fig. 1a for a homogeneous applied field (h_z, h_x), exhibits paramagnetism when applied fields dominate and antiferromagnetism when interactions dominate.

The approach to constructing a magnetic Hamiltonian that we use here was proposed in ref. 31 in the context of experiments reported in ref. 10, in which a gradient was applied to measure the insulating properties of the Mott state. In ref. 31 it was shown that under the influence of such field gradients, the dynamics of a one-dimensional Mott insulator map onto the aforementioned Ising model (see Methods) in the neighbourhood of $(h_z, h_x) = (1,0)$ (Fig. 1b).

In the Mott insulator regime ($U \gg t$), it is energetically forbidden for the atoms to tunnel if the tilt per lattice site, E, differs from the on-site atom–atom interaction U. Hence, the system remains in a state with one atom per lattice site while $E < U$ (Fig. 2a). As the tilt approaches the interaction strength ($E = U$), each atom is free to tunnel onto its neighbour, if that neighbour has not itself tunnelled (Fig. 2b). This nearest-neighbour constraint is the source of the effective spin–spin interaction. If the tilt E is increased sufficiently slowly through the transition to ensure that the system remains near the many-body ground state, density wave ordering results (Fig. 2c).

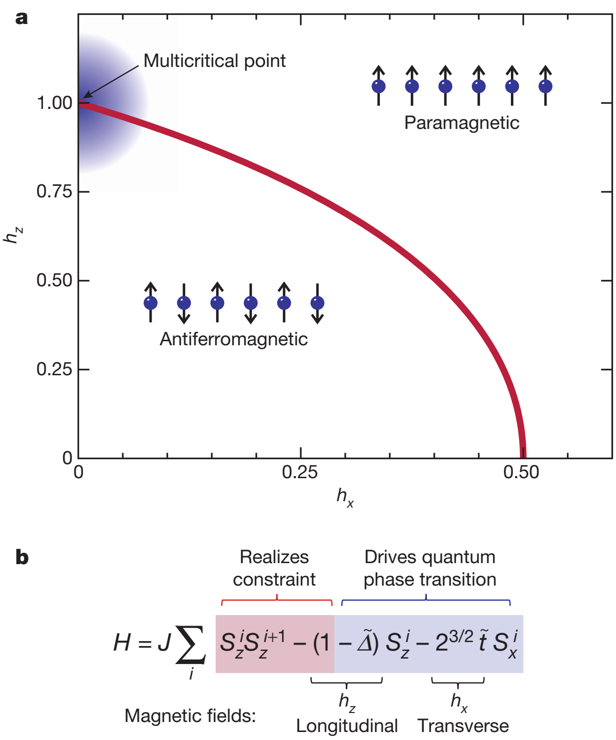


Figure 1 | Spin model and its phase diagram. **a**, An antiferromagnetic one-dimensional Ising chain in longitudinal (h_z) and transverse(h_x) magnetic fields exhibits two phases at zero temperature[39,40]: an antiferromagnetic phase in weak fields, and a paramagnetic phase in strong fields. These phases are separated by a second-order phase transition (red line), except at the multicritical point $(h_z, h_x) = (1, 0)$, where the absence of quantum fluctuations produces a classical first-order transition. The region accessible in our experiment is highlighted in blue. **b**, Here the Hamiltonian may be decomposed into a constraint term (shaded red) that prevents adjacent 'down'-spins, and field terms (shaded blue) that drive the phase transition. See text for details.

The mapping onto a spin-½ model arises as each atom has only two possible positions: an atom that has not tunnelled corresponds to an 'up' spin, and one that has, to a 'down' spin. Figure 2d shows the spin configurations corresponding to various atom distributions in the optical lattice. The transition from a uniform phase at small tilt to a density wave phase at large tilt then corresponds to a transition from a paramagnetic phase to an antiferromagnetic phase in the spin model. The longitudinal field h_z thus arises from the lattice tilt, and the transverse field h_x from tunnelling. The mapping between Bose–Hubbard and spin models (see Methods) is given by $(h_z, h_x) = (1-\tilde{\Delta}, 2^{3/2}\tilde{t})$, $\tilde{t} = t/J$, $\tilde{\Delta} = \Delta/J = (E-U)/J$, with $t = 10(4)$ Hz the single-particle tunnelling rate, $J \approx U = 413(19)$ Hz the constraint term, Δ the energy cost to tunnel, and $\tilde{t}$ and $\tilde{\Delta}$ the normalized tunnelling rate and tunnelling energy cost. Numbers in parentheses are 1σ uncertainties in measured quantities.

Spatially inhomogeneous tilts in the optical lattice can produce site-to-site variation of h_z, which affects the critical behaviour by inducing different sites to undergo the transition at different applied tilts. The resulting many-body energy gaps, dynamical timescales[41,42], and entropy of entanglement[43] then differ from the homogeneous case. Controlling such inhomogeneities is thus crucial for studies of magnetism.

Higher order processes can produce atom configurations that are not within the spin model (see Supplementary Information). We can thus study ground-state dynamics and low-energy excitations of the equivalent Ising model, but not high-energy excitations associated with adjacent 'down' spins. These restrictions admit investigation of the Ising physics only in the neighbourhood of the multicritical point $(h_z, h_x) \approx (1, 0)$. This regime is of particular theoretical interest because the model is here not exactly solvable[40]. It is nonetheless in the Ising universality class[31], and so a study of its critical physics would provide insight into the behaviours of the more commonly considered transverse ($h_z = 0$) Ising model.

Extracting spin observables

We detect magnetic ordering with single-site resolution by using our quantum gas microscope. The microscope is sensitive only to the parity of the site occupation number[14], and so paramagnetic domains (with one atom per lattice site) should appear as entirely bright regions, and antiferromagnetic domains (with alternating 0–2–0–2 occupation) as entirely dark regions. Within the resonant Hubbard subspace which maps to the spin model, the mean z-projection of spin may be related to odd-occupation probability at site i, p^i_{odd}, according to (see Methods):

$$\overline{\langle S_z^i \rangle} = \frac{1}{2}\overline{p^i_{\mathrm{odd}}}$$

Here angle brackets denote realization-averages, and overbars region-averages.

Such an occupation measurement enables us to locally identify magnetic domains and estimate their size, but is not a direct measurement of the antiferromagnetic order parameter[31], and does not reflect the broken symmetry in the antiferromagnetic phase (see Methods). Accordingly, we also study the antiferromagnetic order parameter more directly through one-dimensional quantum noise interferometry[44].

Observing the phase transition

Our experiments begin with a Mott insulator of ^{87}Rb atoms in a two-dimensional optical lattice (with spacing $a = 680$ nm and a depth of $35E_r$, where E_r is the lattice recoil energy) in the focal plane of a high-resolution imaging system; this imaging system allows detection of single atoms on individual lattice sites, as described in previous work[14,15]. The value of E_r is given by $h^2/8ma^2$, where h is Planck's constant and m the mass of ^{87}Rb. We generate our effective h_z by tilting the lattice potential by E per lattice site, achieved by a magnetic

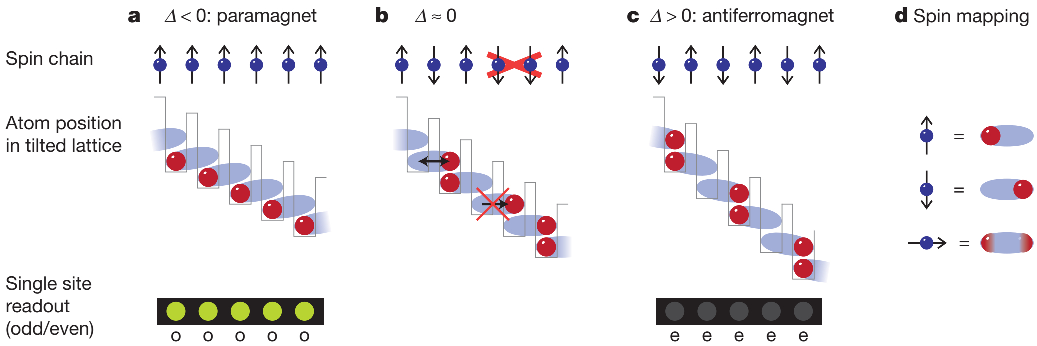


Figure 2 | Tilted Hubbard model and mapping to spin model. Consider first the middle row. **a**, When a Mott insulator is tilted by E per lattice site, it maintains unity occupancy until E reaches the on-site interaction energy U. **b**, As the energy cost to tunnel, $\Delta = E - U$, vanishes, an atom can tunnel to the neighbouring site if the atom on that site has not itself tunnelled. Otherwise, the tilt E inhibits tunnelling, producing a strong constraint. **c**, Tilting further, the system undergoes a transition into a doubly degenerate staggered phase. **d**, Here we show how this system maps onto interacting spin-½ particles, whose two spin states correspond to the two possible locations of each atom. The tunnelling constraint forbids adjacent down spins, realizing a spin–spin interaction. Top row: the initial Mott insulator corresponds to a paramagnet (**a**), the state at resonant tilt to an entangled (critical) spin configuration (**b**), and staggered ordering at large tilt to an antiferromagnet (**c**). Bottom row: parity-sensitive site-resolved imaging results in bright paramagnetic (**a**; o, odd), and dark antiferromagnetic (**c**; e, even) domains.

field gradient along the x direction (defined in Fig. 3A). We then prepare decoupled one-dimensional chains and null the harmonic confinement (see Methods). Finally we ramp (that is, increase) the lattice tilt, tuning the system across the transition at a fixed rate, and stop at various times to observe spin ordering.

We initiate the gradient ramp on the paramagnetic side of the phase transition (typically at $E/U = 0.7$), as the initial Mott state has good overlap with the paramagnetic ground state (Fig. 3A, a). At the end of the ramp ($E/U = 1.2$), we observe even occupation with probability 0.90(2) (Fig. 3A, b), as expected for an antiferromagnetic phase where the spin–spin interaction overwhelms the effective field h_z. In between, density-wave (antiferromagnetic) ordered regions begin to form, as shown in Fig. 3A, c. Figure 3C shows p_{odd} (averaged over 40 sites and 6 realizations) at various times during this ramp. A crucial characteristic of an adiabatic transition is that it is reversible. Figure 3B shows p_{odd} during a ramp from a paramagnetic phase to an antiferromagnetic phase and back in 500 ms. The recovery of the singly occupied sites is evidence of the reversibility of the process, and hence that the state after the forward ramp is an antiferromagnet.

We directly verify the existence of staggered ordering in the antiferromagnetic phase via a quantum noise correlation measurement[44] after a one-dimensional expansion (see Methods). The resulting one-dimensional spatial autocorrelation is plotted in Fig. 3D at both the beginning (Fig. 3D, a) and end (Fig. 3D, b) of the ramp from the paramagnetic phase to the antiferromagnetic phase. In the paramagnetic phase, the spectrum exhibits peaks at momentum difference $P = h/a$, characteristic of a Mott insulator[45]. In the antiferromagnetic phase, peaks appear at $P = h/2a$, indicating spatial ordering with twice the wavelength.

Single-site study of the transition

A high-resolution study reveals that in the presence of harmonic confinement the spins undergo the transition sequentially, owing to the resulting spatial variation of the effective longitudinal field. Figure 4a shows p_{odd} versus tilt for two rows of a harmonically confined Mott insulator, separated by seven lattice sites. That these rows tune through resonance at different tilts can be understood from the energy level diagram (Fig. 4b). We realize a homogeneous field Ising model by eliminating the harmonic confinement (see Methods) immediately before the slow ramp into the antiferromagnetic phase. The homogeneity is now only limited by residual lattice beam disorder, and accordingly, Fig. 4c demonstrates that different rows undergo the transition almost simultaneously, as anticipated theoretically (Fig. 4d).

After nulling the harmonic confinement, we use high-resolution imaging to study the transition on the single-site level. This allows us

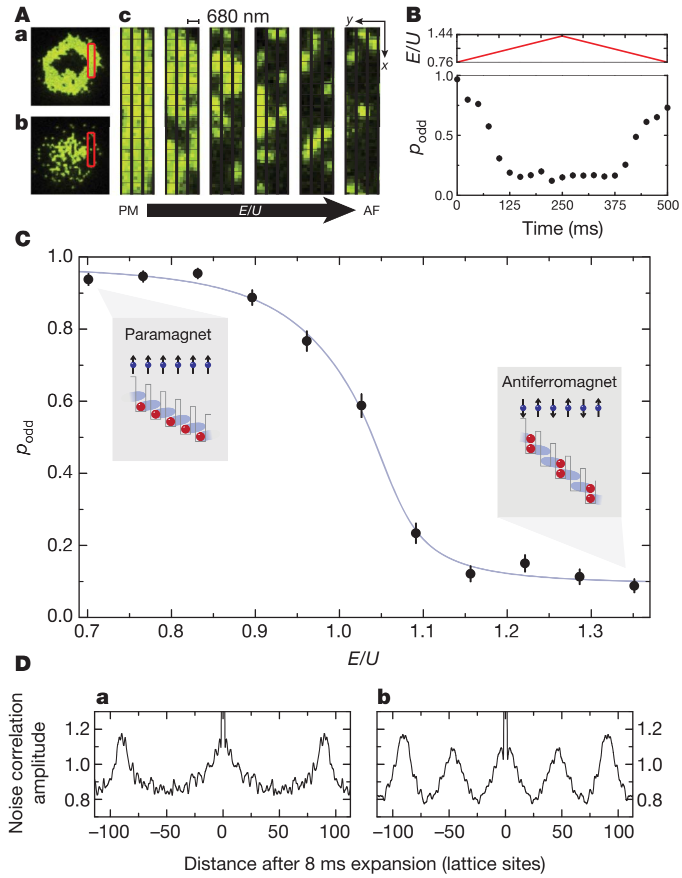


Figure 3 | Probing the paramagnet to antiferromagnet phase transition. **A**, Single-shot images as the tilt is swept through the phase transition in 250 ms. **A**, **a**, $n = 1$ (bright) and $n = 2$ (centre dark) Mott insulator shells in the paramagnetic (PM) phase. **A**, **b**, 'Inverted' shells characteristic of the staggered ordering of the antiferromagnetic (AF) phase. **A**, **c**, Several chains (red rectangles in **a** and **b**) of the $n = 1$ shell at various times during the sweep (left to right, in ms: 0, 50, 100, 150, 175, 250), showing antiferromagnetic domain formation. **B**, Demonstration of reversibility by tuning across the transition and back. Data points are p_{odd}, red line is tilt E/U, both versus time. **C**, A closer look at the paramagnetic to antiferromagnetic quantum phase transition in an $n = 1$ shell. Error bars, 1σ statistical errors in the region-averaged mean p_{odd}. Blue curve is a guide for the eye. Insets, state of the system at the beginning and end of the sweep. **D**, One-dimensional noise correlation measurement, with peaks at $P = h/a$ in the paramagnetic phase (**a**), and additional peaks at $P = h/2a$ in the antiferromagnetic phase (**b**).

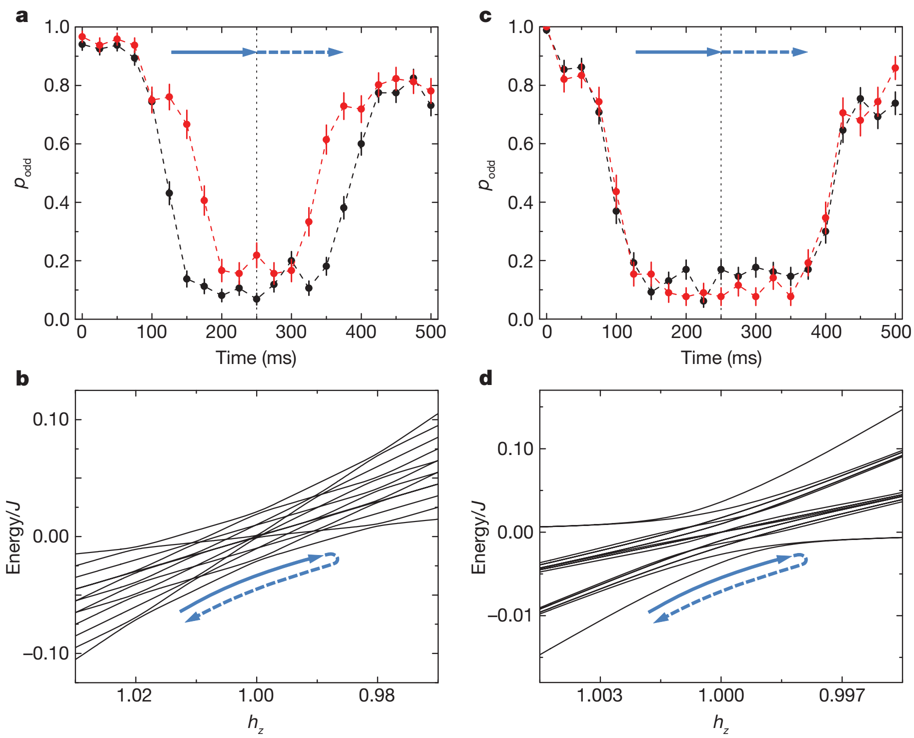


Figure 4 | Effect of harmonic confinement. a, For a ramp (in E/U; see Fig. 3B) across the transition and back in 500 ms, harmonic confinement broadens the transition region, inducing two rows (red and black data points) separated by seven lattice sites to go dark at different applied tilts. Red and black dashed lines are guides to the eye. **b**, Energy spectrum (black lines) of a one-dimensional Ising chain in a longitudinal field gradient (see Supplementary Information). Avoided crossings of the ground state correspond to single spin-flips. **c**, After nulling the confinement, the two rows undergo the transition together. **d**, Energy spectrum for a homogeneous chain. A single avoided crossing drives the many-body transition, whose gap scales inversely with system size. Solid and dashed blue arrows denote forward- and subsequent reverse-tilt ramps in all panels. Error bars are 1σ statistical uncertainties in the region-averaged mean of p_{odd}.

to focus on a single six-site chain with particularly low inhomogeneity, which we will study for the remainder of this Article. We identify such a chain by imaging individual lattice sites as the system is tuned across the paramagnetic–antiferromagnetic transition. Figure 5 shows the average occupation (over 43 realizations) of each of the six adjacent sites (black curves), versus tilt, as we ramp E/U across the transition. The r.m.s. variation in the fitted centres of the single-site transition curves (see Supplementary Fig. 1) is 6 Hz. This variation is significantly less than their mean 10–90% width of 105(30) Hz which corresponds to the effective transverse field $2^{3/2}t = 28(9)$ Hz. By quickly jumping across the transition with tunnelling inhibited, and then ramping slowly across the transition in reverse with tunnelling allowed (red curves in Fig. 5, taken under slightly different conditions), we are able to rule out large, localized potential steps that would otherwise prevent individual spins from flipping. The curves in Fig. 5 provide our best estimate of the inhomogeneity. However, exact determination of the site-to-site disorder using this technique is complicated by the many-body nature of the observed transition. New techniques, such as site-resolved modulation spectroscopy, would be necessary to ensure sufficient homogeneity for studies of criticality in long, uniform Ising chains.

Domain formation in a one-dimensional Ising chain

As the system is tuned through the transition by ramping E/U, quantum fluctuations induce the formation of antiferromagnetic domains, which appear as uninterrupted strings of dark lattice sites. Figure 6a shows the observed mean length-weighted dark domain length extracted from 43 single-shot images per tilt, as the system is tuned from the paramagnetic phase into the antiferromagnetic phase. The dark domain length is here defined as the number of contiguous dark sites in the previously identified homogeneous six-site chain (see Supplementary Information). The mean dark domain length reaches

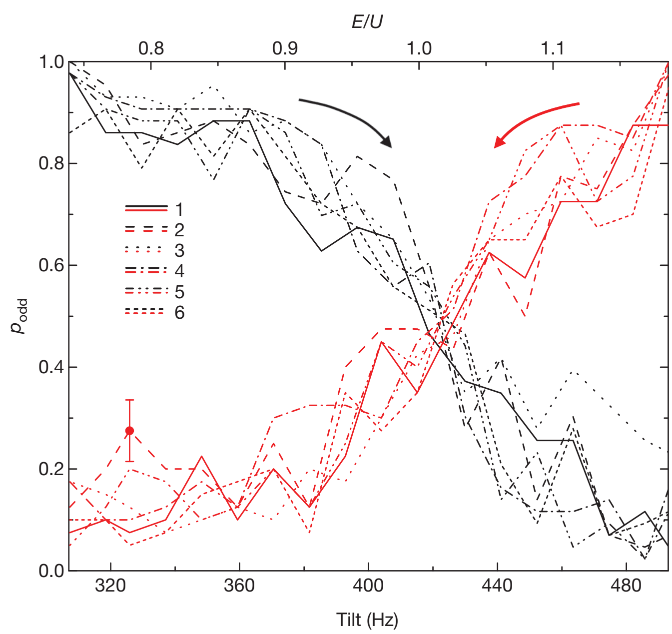


Figure 5 | Site-resolved transition in near-homogeneous Ising model. Shown is p_{odd} for six individual lattice sites forming a contiguous one-dimensional chain, plotted against the tilt, for both forward (black) and reverse (red) ramps of 250 ms. The spins are observed to undergo the transition at the same applied field to within the curve width, set by quantum fluctuations. Typical 1σ statistical error bar is shown. The reverse curve demonstrates the preparation of the highest-energy state of the restricted spin Hamiltonian, exhibiting paramagnetic ordering on the antiferromagnetic side of the transition, and antiferromagnetic ordering on the paramagnetic side.

4.9(2) sites, giving evidence that the average antiferromagnetic domain size approaches the system size.

We next investigate the effect of ramp rate on the transition from the paramagnetic phase into the antiferromagnetic phase in the homogeneous chain. The blue points in Fig. 6b show p_{odd} as a function of ramp speed across the transition, indicating qualitatively the fraction of the system that has not undergone transition into antiferromagnetic domains of any size. The time required to flip the spins is ~50 ms, consistent with tunnelling-induced quantum fluctuations driving the transition. The black points in Fig. 6b are the mean length-weighted dark domain length as a function of ramp rate, and approach the system size of six sites. The remaining defects probably result from heating during the ramp, and from initial defects of the Mott insulator (see Supplementary Information), as well as imperfect overlap of the initial Mott insulator with the paramagnetic state, which at initial tilt Δ_{i} is $1-O(t^2/\Delta_{\mathrm{i}}^2)$.

Whereas the antiferromagnetic domain formation discussed thus far occurs in a system near its many-body ground state, we can also produce antiferromagnetic domains corresponding to the highest-energy state of the constrained Hilbert space. This is achieved by starting with the Mott insulator and rapidly ramping the field gradient through the transition point with tunnelling inhibited, then adiabatically ramping back with tunnelling permitted. This prepares a paramagnet on the antiferromagnetic side of the transition, then adiabatically converts it into an antiferromagnet on the paramagnetic side. The resulting data are shown in red in Figs 5 and 6a, demonstrating that these high-energy states are sufficiently long-lived to support domain formation. Similar ideas have been proposed for preparation of difficult-to-access many-body states using the highest energy states of Hamiltonians with easily prepared ground states[46].

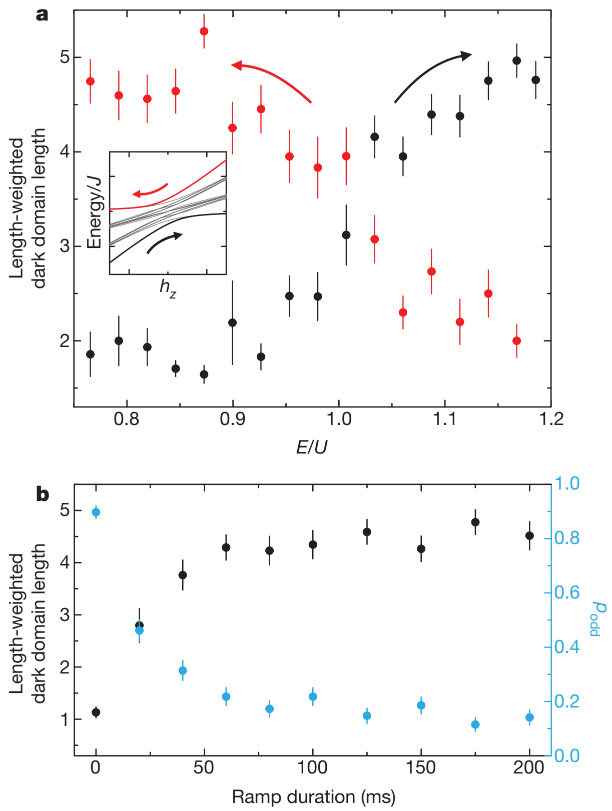


Figure 6 | Dynamics of antiferromagnetic domain formation. **a**, Within a single six-site chain with low disorder, the mean dark-domain length is plotted against tilt for forward (black) and reverse (red) ramps, as it grows to nearly the chain length. The reverse ramp produces antiferromagnetic domains on the paramagnetic side of the transition—corresponding to the highest-energy states of the spin Hamiltonian (inset). Arrows denote ramp direction. **b**, Within the same chain, p_{odd} (blue) and dark domain length (black) are plotted against ramp time from $E/U=0.7$ to $E/U=1.2$. Error bars for dark domain lengths are 1σ statistical uncertainties arising from the number of detected domains of each length; those for p_{odd} are 1σ statistical uncertainties in the mean.

Conclusions and outlook

We have experimentally realized a quantum simulation of an Ising chain in the presence of longitudinal and transverse fields. By varying the applied longitudinal field, we drive a transition between paramagnetic and antiferromagnetic phases, and verify the formation of spin domains via both direct *in situ* imaging and noise correlation in expansion. We study the adiabaticity requirements for transition dynamics, observing a timescale consistent with tunnelling-induced quantum fluctuations. By tuning rapidly through the transition and ramping back across it slowly, we prepare the highest-energy state of a many-body spin Hamiltonian.

We introduce and implement a novel route to studying low-entropy magnetism in optical lattices. Complexities associated with the cooling of spin mixtures[35–37] are circumvented by using a low-entropy, gapped Mott insulator as an initially spin-polarized state, and then adiabatically opening a spin degree of freedom[38]. The achieved entropies are so low that the residual tilt inhomogeneity, and not entropy, limits domain size. This recipe is directly applicable to more traditional approaches to quantum magnetism, including those using super-exchange interactions[33].

The spin-mapping demonstrated here realizes strong effective magnetic interactions, and opens new possibilities for studies of quantum magnetism and criticality. It will be interesting to investigate the effect of various types of controlled disorder on criticality and transition dynamics, as well as the possible existence of non-thermalizing states[47]. A particularly intriguing direction is the extension of the tilted Mott insulator physics to other models[48], and to higher dimensions. The square lattice geometry generates phases with longitudinal density wave ordering and transverse superfluidity[31]. More sophisticated geometries can induce frustration, resulting in novel phases such as quantum-liquids and valence-bond solids[49].

METHODS SUMMARY

We begin with a single-layer two-dimensional Mott insulator of ^{87}Rb atoms in a $35E_{\mathrm{r}}$ lattice with 680-nm spacing as described in previous work. A magnetic field gradient along the x direction is ramped up to tilt the lattice. The lattice depth is then ramped to inhibit tunnelling transverse to the tilt, and to enhance tunnelling along the tilt. At the same time, the optical potential providing harmonic confinement is ramped down. The gradient is then ramped adiabatically through the transition point using a linear ramp. We can then perform either *in situ* imaging or a one-dimensional expansion of the chains to achieve noise correlation interferometry. In both cases, we use fluorescence imaging after pinning the atoms in a deep lattice to obtain the density distribution with single atom/single-lattice-site resolution. Fluorescence detection gives the atom number modulo 2 owing to light-assisted collisions of atoms on each lattice site during imaging.

Lattice depths are calibrated using Kapitza–Dirac scattering to 15%, however the width of single-site transition regions was found to be a more sensitive probe of the longitudinal tunnelling rate and hence of the longitudinal lattice depth (see Supplementary Fig. 1). The magnetic field gradient is calibrated using lattice modulation spectroscopy (see Supplementary Fig. 2).

We follow ref. 31 in mapping a one-dimensional Mott insulator of spinless bosons in a tilted lattice onto a chain of interacting dipoles (doublon-hole pairs, in a singly occupied Mott shell), and then onto a chain of spin-½ particles with antiferromagnetic Ising interactions in longitudinal and transverse fields.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 January; accepted 15 March 2011.
Published online 13 April 2011.

1. Balents, L. Spin liquids in frustrated magnets. *Nature* **464,** 199–208 (2010).
2. Binder, K. & Young, A. P. Spin glasses: experimental facts, theoretical concepts, and open questions. *Rev. Mod. Phys.* **58,** 801–976 (1986).

3. Kitaev, A. Anyons in an exactly solved model and beyond. *Ann. Phys.* **321,** 2–111 (2006).
4. Sachdev, S. *Quantum Phase Transitions* (Cambridge Univ. Press, 2001).
5. Anderson, P. W. The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235,** 1196–1198 (1987).
6. Rüegg, C. *et al.* Quantum magnets under pressure: controlling elementary excitations in $TlCuCl_3$. *Phys. Rev. Lett.* **100,** 205701 (2008).
7. Coldea, R. *et al.* Quantum criticality in an Ising chain: experimental evidence for emergent E8 symmetry. *Science* **327,** 177–180 (2010).
8. Lewenstein, M. *et al.* Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond. *Adv. Phys.* **56,** 243–379 (2007).
9. Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80,** 885–964 (2008).
10. Greiner, M., Mandel, O., Esslinger, E., Haensch, T. W. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415,** 39–44 (2002).
11. Sachdev, S. Quantum magnetism and criticality. *Nature Phys.* **4,** 173–185 (2008).
12. Zhang, X., Hung, C.-L., Tung, S.-K., Gemelke, N. & Chin, C. Exploring quantum criticality based on ultracold atoms in optical lattices. Preprint at ⟨http://arXiv.org/abs/1101.0284⟩ (2010).
13. Hung, C.-L., Zhang, X., Gemelke, N. & Chin, C. Observation of scale invariance and universality in two-dimensional Bose gases. *Nature* **470,** 236–239 (2011).
14. Bakr, W. S., Gillen, J. I., Peng, A., Foelling, S. & Greiner, M. A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice. *Nature* **462,** 74–77 (2009).
15. Bakr, W. S. *et al.* Probing the superfluid-to-Mott insulator transition at the single-atom level. *Science* **329,** 547–550 (2010).
16. Sherson, J. F. *et al.* Single-atom-resolved fluorescence imaging of an atomic Mott insulator. *Nature* **467,** 68–72 (2010).
17. Weitenberg, C. *et al.* Single-spin addressing in an atomic Mott insulator. *Nature* **471,** 319–324 (2011).
18. Jo, G. *et al.* Itinerant ferromagnetism in a Fermi gas of ultracold atoms. *Science* **325,** 1521–1524 (2009).
19. Friedenauer, A., Schmitz, H., Glueckert, J. T., Porras, D. & Schaetz, T. Simulating a quantum magnet with trapped ions. *Nature Phys.* **4,** 757–761 (2008).
20. Kim, K. *et al.* Quantum simulation of frustrated Ising spins with trapped ions. *Nature* **465,** 590–593 (2010).
21. Ni, K.-K. *et al.* A high phase-space-density gas of polar molecules. *Science* **322,** 231–235 (2008).
22. Saffman, M. & Walker, T. G. Quantum information with Rydberg atoms. *Rev. Mod. Phys.* **82,** 2313–2363 (2010).
23. Lukin, M. D. *et al.* Dipole blockade and quantum information processing in mesoscopic atomic ensembles. *Phys. Rev. Lett.* **87,** 037901 (2001).
24. Büchler, H. P. *et al.* Strongly correlated 2D quantum phases with cold polar molecules: controlling the shape of the interaction potential. *Phys. Rev. Lett.* **98,** 060404 (2007).
25. Weimer, H., Müller, M. & Lesanovsky, I. A Rydberg quantum simulator. *Nature Phys.* **6,** 382–388 (2010).
26. Lee, P. J. *et al.* Sublattice addressing and spin-dependent motion of atoms in a double-well lattice. *Phys. Rev. Lett.* **99,** 020402 (2007).
27. Soltan-Panahi, P. *et al.* Multi-component quantum gases in spin-dependent hexagonal lattices. *Nature Phys.* advance online publication doi:10.1038/nphys1916 (13 February 2011).
28. Fölling, S. *et al.* Direct observation of second-order atom tunnelling. *Nature* **448,** 1029–1032 (2007).
29. Fisher, M. P. A., Weichman, P. B., Grinstein, G. & Fisher, D. S. Boson localization and the superfluid-insulator transition. *Phys. Rev. B* **40,** 546–570 (1989).
30. Jaksch, D., Bruder, C., Cirac, J. I., Gardiner, C. W. & Zoller, P. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.* **81,** 3108–3111 (1998).
31. Sachdev, S., Sengupta, K. & Girvin, S. M. Mott insulators in strong electric fields. *Phys. Rev. B* **66,** 075128 (2002).
32. Duan, L. M., Demler, E. & Lukin, M. D. Controlling spin exchange interactions of ultracold atoms in optical lattices. *Phys. Rev. Lett.* **91,** 090402 (2003).
33. Trotzky, S. *et al.* Time-resolved observation and control of superexchange interactions with ultracold atoms in optical lattices. *Science* **319,** 295–299 (2008).
34. Capogrosso-Sansone, B., Soeyler, S. G., Prokof'ev, N. V. & Svistunov, B. V. Critical entropies for magnetic ordering in bosonic mixtures on a lattice. *Phys. Rev. A* **81,** 053622 (2010).
35. Weld, D. M. *et al.* Spin gradient thermometry for ultracold atoms in optical lattices. *Phys. Rev. Lett.* **103,** 245301 (2009).
36. Medley, P., Weld, D., Miyake, H., Pritchard, D. E. & Ketterle, W. Spin gradient demagnetization cooling of ultracold atoms. Preprint at ⟨http://arxiv.org/abs/1006.4674⟩ (2010).
37. McKay, D. & DeMarco, B. Cooling in strongly correlated optical lattices: prospects and challenges. Preprint at ⟨http://arxiv.org/abs/1010.0198⟩ (2010).
38. García-Ripoll, J. J., Martin-Delgado, M. A. & Cirac, J. I. Implementation of spin Hamiltonians in optical lattices. *Phys. Rev. Lett.* **93,** 250405 (2004).
39. Novotny, M. A. & Landau, D. P. Zero temperature phase diagram for the $d=1$ quantum Ising antiferromagnet. *J. Magn. Magn. Mater.* **54-57,** 685–686 (1986).
40. Ovchinnikov, A. A., Dmitriev, D. V., Krivnov, V. Y. & Cheranovskii, V. O. Antiferromagnetic Ising chain in a mixed transverse and longitudinal magnetic field. *Phys. Rev. B* **68,** 214406 (2003).
41. Imry, Y. & Ma, S.-k. Random-field instability of the ordered state of continuous symmetry. *Phys. Rev. Lett.* **35,** 1399–1401 (1975).
42. Dziarmaga, J. Dynamics of a quantum phase transition in the random Ising model: logarithmic dependence of the defect density on the transition rate. *Phys. Rev. B* **74,** 064416 (2006).
43. Vidal, G., Latorre, J. I., Rico, E. & Kitaev, A. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.* **90,** 227902 (2003).
44. Altman, E., Demler, E. & Lukin, M. D. Probing many-body states of ultracold atoms via noise correlations. *Phys. Rev. A* **70,** 013603 (2004).
45. Fölling, S. *et al.* Spatial quantum noise interferometry in expanding ultracold atom clouds. *Nature* **434,** 481–484 (2005).
46. Sørensen, A. S. *et al.* Adiabatic preparation of many-body states in optical lattices. *Phys. Rev. A* **81,** 061603 (2010).
47. Bañuls, M. C., Cirac, J. I. & Hastings, M. B. Strong and weak thermalization of infinite nonintegrable quantum systems. *Phys. Rev. Lett.* **106,** 050405 (2011).
48. Plötz, P., Schlagheck, P. & Wimberger, S. Effective spin model for interband transport in a Wannier-Stark lattice system. *Eur. Phys. J. D.* doi:10.1140/epjd/e2010-10554-7 (2010).
49. Pielawa, S., Kitagawa, T., Berg, E. & Sachdev, S. Correlated phases of bosons in tilted, frustrated lattices. Preprint at ⟨http://arxiv.org/abs/1101.2897⟩ (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Demler, W. Ketterle, T. Kitagawa, M.D. Lukin, S. Pielawa and S. Sachdev for discussions. This work was supported by the Army Research Office DARPA OLE programme, an AFOSR MURI programme, and by grants from the NSF.

Author Contributions All authors contributed to the construction of the experiment, the collection and analysis of the data, and the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.G. (greiner@physics.harvard.edu).

METHODS

Mapping onto the spin model. We follow ref. 31 in formally mapping a one-dimensional Mott insulator of spinless bosons in a tilted lattice onto a chain of interacting dipoles (doublon-hole pairs, in a singly occupied Mott shell), and then onto a chain of spin-½ particles with antiferromagnetic Ising interactions in longitudinal and transverse fields. In a homogeneously tilted lattice, the one-dimensional Bose–Hubbard Hamiltonian reads:

$$H=-t\sum_j\left(a_j^\dagger a_{j+1}+a_j a_{j+1}^\dagger\right)+\frac{U}{2}\sum_j n_j(n_j-1)-E\sum_j jn_j$$

Here t is the nearest-neighbour tunnelling rate, U is the on-site interaction, E is the tilt per lattice site, $a_j^\dagger(a_j)$ is the creation (annihilation) operator for a particle on site j, and $n_j=a_j^\dagger a_j$ is the occupation number operator on site j.

For a tilt near $U=E$, the on-site interaction energy cost for an atom to tunnel onto its neighbour is almost precisely cancelled by the tilt energy. If one starts in a Mott insulator with M atoms per site, an atom can then resonantly tunnel onto the neighbouring site to produce a dipole excitation, with a pair of sites with $M+1$ and $M-1$ atoms. The resonance condition is only met if adjacent sites contain equal numbers of atoms, so only one dipole can be created per link and neighbouring links cannot both support dipoles. We define a dipole creation operator $d_j^\dagger=\frac{a_j a_{j+1}^\dagger}{\sqrt{M(M+1)}}$

The Bose–Hubbard Hamiltonian above can hence be mapped onto the dipole Hamiltonian:

$$H=-\sqrt{M(M+1)}t\sum_j\left(d_j^\dagger+d_j\right)+(U-E)\sum_j d_j^\dagger d_j$$

subject to the constraints $d_j^\dagger d_j\leq 1, d_{j+1}^\dagger d_{j+1} d_j^\dagger d_j=0$

The factor of $\sqrt{M(M+1)}$ arises owing to bosonic enhancement.

To map from the dipole Hamiltonian to the spin-½ Hamiltonian, we define a link without (with) a dipole excitation to be an up (down) spin along $\hat{z}$. Then the creation/annihilation of dipoles are related to the flipping of spins, and we can write:

$$S_z^j=\tfrac{1}{2}-d_j^\dagger d_j,\ S_x^j=\tfrac{1}{2}\left(d_j^\dagger+d_j\right),\ \text{and}\ S_y^j=\tfrac{i}{2}\left(d_j^\dagger-d_j\right)$$

The constraint forbidding adjacent dipoles can be implemented by introducing a positive energy term $Jd_{j+1}^\dagger d_{j+1}d_j^\dagger d_j=J\left(S_z^{j+1}-\frac{1}{2}\right)\left(S_z^j-\frac{1}{2}\right)$ to the Hamiltonian, where J is of order U. This term gives rise to nearest-neighbour interactions and an effective longitudinal field for the spins.

Defining $\Delta=E-U$, the Hamiltonian for the spins now reads:

$$H=J\sum_j S_z^j S_z^{j+1}-2\sqrt{M(M+1)}t\sum_j S_x^j-(J-\Delta)\sum_j S_z^j$$

$$=J\sum_j\left(S_z^j S_z^{j+1}-h_x S_x^j-h_z S_z^j\right)$$

The dimensionless fields are defined as $h_x=2^{3/2}t/J=2^{3/2}\tilde{t}, h_z=\left(1-\frac{\Delta}{J}\right)=1-\tilde{\Delta}$, with M set to one as in our experiment.

Experimental details. Our experiments start with a single-layer two-dimensional Mott insulator of ^{87}Rb atoms in a $35E_r$ lattice with 680-nm spacing as described in previous work. The atoms are in the $|F=1, m_F=-1\rangle$ state and the initial fidelity of the Mott insulator is 0.95(2), with local fidelities as high as 0.98. A magnetic field gradient along the x direction is ramped up within 8 ms to tilt the lattice potential by $0.7U$ per lattice site, just below the transition point[31,40] at $E=U+1.85t$ ($h_z=1-0.66h_x$). At this point, the depth of the lattice along the chains is ramped down to $14E_r$, while the lattice transverse to the chains is ramped up to $45E_r$, within 2 ms. At the same time, the optical potential providing harmonic confinement is ramped down. Tunnelling between chains is negligible over the experimental timescale (see Supplementary Information for further discussion). The gradient is then ramped adiabatically through the transition point using a linear ramp that ends at a tilt of $1.2U$ per lattice site, typically within 250 ms.

We can then perform either an *in situ* measurement or a one-dimensional expansion of the chains to achieve noise correlation interferometry. In both cases, we use fluorescence imaging after pinning the atoms in a deep lattice to obtain the density distribution with single-atom/single-lattice-site resolution. Images far on the Mott side of the transition are used to select chains of atoms within the first shell of the insulator. The phase transition is then studied only within these chains, with quantitative curves using data only from the single chain with lowest disorder.

For noise correlation measurements, we first increase the lattice depth along the chains to $35E_r$ within 8 ms and then rapidly switch off that lattice to realize a one-dimensional expansion. Simultaneously with the lattice turn-off, the magnetic field gradient and the lattice along the chain are switched off, while the interchain lattice and the potential confining the atoms in the third direction remain on. After a one-dimensional expansion for 8 ms, the atoms are pinned for imaging. To extract information about density wave ordering in the chains, several hundred images (100 for paramagnetic phase, 300 for antiferromagnetic phase), each containing 15 chains, are fitted to extract the atom positions, and then spatially autocorrelated and averaged as described in ref. 45. In principle, the mean domain size can be extracted from the $P=h/2a$ peak width, however our peaks are broadened by finite expansion time and interactions during expansion, as well as aberration arising from the fact that the one-dimensional expansion is performed not in free space but in slightly corrugated confining tubes. Consequently domain size information is difficult to extract.

Lattice depths are calibrated to 15% using Kapitza–Dirac scattering, however the width of single-site transition regions was found to be a more sensitive probe of the longitudinal tunnelling rate and hence of the longitudinal lattice depth (see Supplementary Fig. 1), and accordingly was used throughout this Article.

The magnetic field gradient is calibrated using lattice modulation spectroscopy. In the presence of a potential gradient E per lattice site, modulation of the lattice depth along the chains causes resonant excitation at two frequencies, $U+E$ and $U-E$, corresponding to an atom in the Mott insulator moving up or down gradient. We detect these excitations as a reduction in the value of p_{odd} using *in situ* imaging (see Supplementary Fig. 2). Using the mean of the two resonances, we obtain the interaction energy $U=430(20)$ Hz at $16E_r$ longitudinal lattice, $45E_r$ transverse lattice (corresponding to $U=413(19)$ Hz at $14E_r$ longitudinal lattice, where the experiment operates, which agrees with a band-structure calculation of 401(25) Hz). The separation between the resonances as a function of applied gradient is used to calibrate E. At zero applied magnetic field gradient, we find the stray gradients to be less than $0.02U$.

Local and long-range observables. *In situ* detection gives the atom number modulo 2, owing to light-assisted collisions of atoms on each lattice site during imaging. In the spin language, the detection parity operator P_i measures the spin–spin correlation between adjacent spins:

$$P_i=4S_z^{i-1}S_z^i$$

The probability that site i has odd occupation is given by $p_{\mathrm{odd}}^i=\frac{1}{2}(1+\langle P_i\rangle)$, where angle brackets denote realization averages. This is then related to the spin observables in the effective model by $\left\langle S_z^{j-1}S_z^j\right\rangle=\frac{1}{2}\left(p_{\mathrm{odd}}^j-\frac{1}{2}\right)$. We average over all the atoms in a region to obtain $p_{\mathrm{odd}}=\overline{p_{\mathrm{odd}}^j}$, which, in combination with the constraint that neighbouring down spins are not allowed, enables us to relate the region-averaged mean z-projection of spin to p_{odd} according to: $\overline{\left\langle S_z^j\right\rangle}=\frac{p_{\mathrm{odd}}}{2}$. This quantity varies smoothly across the transition and depends only weakly on the chain length. On the other hand, the order parameter for the transition $O=\left\langle\left(\frac{1}{N}\sum_j(-1)^j S_z^j\right)^2\right\rangle$ depends on the chain length (Supplementary Fig. 3) and is non-analytic across the transition for a thermodynamic system.

The amplitude of the noise correlation signal at separation d for a chain consisting of a large number of atoms is $C(d)=1+\frac{1}{N^2}\left|\sum_j e^{-i\frac{madj}{\hbar t}}n_j\right|^2$ where N is the number of lattice sites, j is the lattice site index, a is the lattice spacing, m is the mass of the atom, t is the expansion time, n_j is the occupation of the jth site, and $\hbar$ is $h/2\pi$. It can be shown that the correlation signal at $d=\frac{\pi\hbar t}{ma}$ is related to the order parameter $O=C\left(\frac{\pi\hbar t}{ma}\right)-1$.

The finite chain length in the chains reduces the baseline of the correlation signal from 1 by an amount of the order of $1/N$.

ARTICLE

doi:10.1038/nature09821

Catecholamine receptor polymorphisms affect decision-making in *C. elegans*

Andres Bendesky[1], Makoto Tsunozaki[1], Matthew V. Rockman[2], Leonid Kruglyak[3] & Cornelia I. Bargmann[1]

Innate behaviours are flexible: they change rapidly in response to transient environmental conditions, and are modified slowly by changes in the genome. A classical flexible behaviour is the exploration–exploitation decision, which describes the time at which foraging animals choose to abandon a depleting food supply. We have used quantitative genetic analysis to examine the decision to leave a food patch in *Caenorhabditis elegans*. Here we show that patch-leaving is a multigenic trait regulated in part by naturally occurring non-coding polymorphisms in *tyra-3* (*tyramine receptor 3*), which encodes a G-protein-coupled catecholamine receptor related to vertebrate adrenergic receptors. *tyra-3* acts in sensory neurons that detect environmental cues, suggesting that the internal catecholamines detected by *tyra-3* regulate responses to external conditions. These results indicate that genetic variation and environmental cues converge on common circuits to regulate behaviour, and suggest that catecholamines have an ancient role in regulating behavioural decisions.

Despite abundant evidence for heritability of behavioural traits within and between species, only a few naturally varying traits have been associated with polymorphisms in specific genes[1]. Foraging for food is an ecologically relevant, environmentally regulated behaviour that is suitable for genetic analysis, as it can differ between populations of a species that live in different habitats[2]. An essential foraging decision is the choice between exploiting existing resources and exploring other options that may provide new resources. This decision can be described by Charnov's marginal value theorem, which proposes that the optimal time for an animal to leave a foraging ground occurs when local resource levels fall below the average level in the entire habitat[3]. The marginal value theorem was developed for animals foraging for food in patchy environments, but has analogies with diverse decision-making processes in field biology, cognitive neuroscience and economics[2,4–6].

Studies of patch-leaving behaviour in the nematode *C. elegans* have revealed innate, environmental and experience-dependent factors that affect its foraging decisions. *C. elegans* rarely leaves a dense lawn of high-quality bacterial food[7,8], but more frequently leaves lawns of pathogenic bacteria or lawns that are spiked with chemical repellents[9,10]. Males will leave lawns that do not contain potential mates[11], while hermaphrodites leave lawns when animal density is high[12]. In addition, wild-type strains vary in their propensity to leave bacterial lawns based on a genetic polymorphism that affects the G-protein-coupled neuropeptide receptor NPR-1 (refs 12–14). This *npr-1* polymorphism affects many foraging behaviours; low-activity *npr-1* strains aggregate into social feeding groups, move quickly on food, and have altered responses to oxygen, carbon dioxide and pheromones compared to the N2 laboratory strain[15–20]. The high-activity allele of *npr-1* in N2 arose in the laboratory, probably as an adaptation to laboratory conditions[19], so it is not known whether genetic variation affects *C. elegans* foraging in natural environments.

Natural genetic variation within a species can generate diversity in foraging behaviour, as exemplified by the polymorphic *Drosophila melanogaster foraging* (*for*) gene, which encodes a cyclic guanosine monophosphate (cGMP)-dependent protein kinase[21]. A low-activity allele of *for* is present in *Drosophila* sitter larvae, which move slowly on a food patch; a high-activity allele of *for* is present in rover larvae, which move quickly and disperse rapidly[22]. A *for*-related cGMP-dependent kinase affects foraging in honeybees, ants and nematodes, suggesting that diverse animals share molecular mechanisms for behavioural regulation[22,23].

To gain further insight into the genetics and neurobiology of exploratory behaviour in *C. elegans*, we here use quantitative genetic analysis to examine this behaviour's genetic architecture in wild-type strains, and show that genetic variation in multiple loci, including a catecholamine receptor, interacts with environmental conditions to regulate the exploitation–exploration decision.

Multiple loci affect leaving behaviour

Different wild-type strains of *C. elegans* vary in their tendency to leave or remain on a standardized small lawn of bacterial food (Fig. 1a). For example, adult hermaphrodites from the laboratory strain N2 leave the lawn only once every 100 min, whereas animals from the CB4856 (HW) strain isolated from pineapple fields in Hawaii leave the lawn once every 5–6 min (Fig. 1b, Supplementary Movies 1 and 2). To determine the genetic architecture of this behavioural difference between N2 and HW, we quantified leaving rates in 91 N2-HW recombinant inbred advanced intercross lines (RIAILs)[24]. 58 of the RIAILs had low leaving rates comparable to N2, only 6–10 had high leaving rates comparable to HW, and 23 had intermediate rates (Fig. 1c). The excess of low leaving rates and the continuous behavioural distribution in RIAILs suggest that leaving is a multigenic quantitative trait.

Quantitative trait locus (QTL) analysis of the RIAILs uncovered two regions with significant effects on leaving rates, one on the X chromosome and one on chromosome II (Fig. 1d). The X chromosome QTL overlapped with the location of the polymorphic G-protein-coupled neuropeptide receptor NPR-1, which affects many food-related behaviours[12,15]. The *npr-1* polymorphism has previously

[1]Howard Hughes Medical Institute, Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA. [2]Department of Biology and Center for Genomics and Systems Biology, New York University, New York, New York 10003, USA. [3]Howard Hughes Medical Institute, Lewis-Sigler Institute for Integrative Genomics and Department of Ecology and Evolutionary Biology, Carl Icahn Laboratory, Princeton University, Princeton, New Jersey 08544, USA.

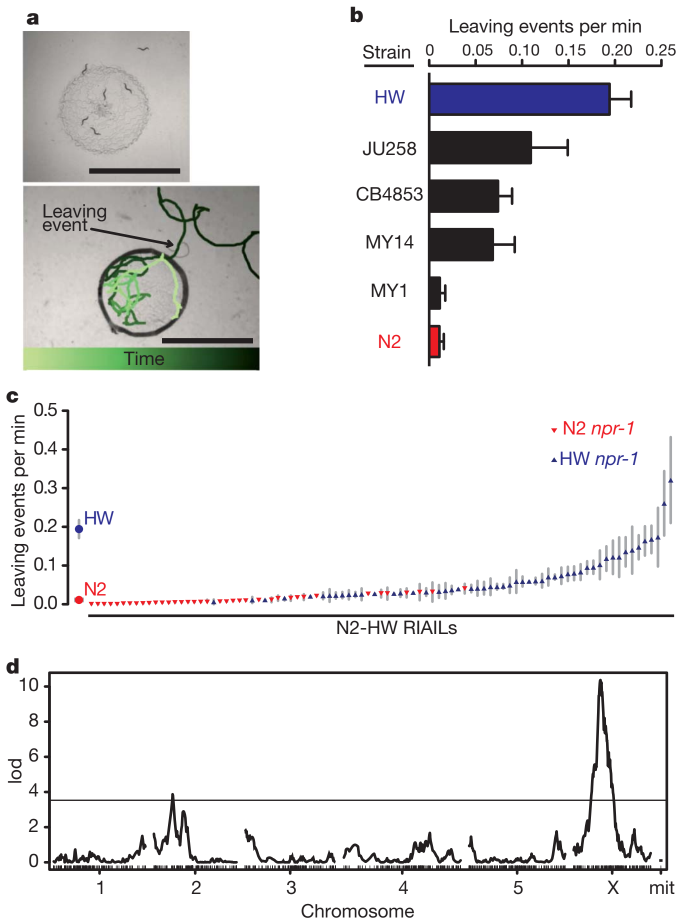


Figure 1 | Lawn-leaving behaviour varies between wild-type *C. elegans* strains. **a**, Lawn-leaving assays. Top, six adult HW hermaphrodites on a bacterial lawn. One animal has left the lawn and one is leaving. Bottom, track of a HW animal during 5 min of an assay. The track is colour-coded to show passage of time (colour scale at bottom). The border of the lawn is outlined. Scale bar, 6 mm. **b,** Leaving rates of six wild-type strains. **c**, Leaving rates of 91 N2-HW recombinant inbred advanced intercross lines (RIAILs)[24] and parental strains. **d**, QTL analysis of RIAILs shown in **c**. The horizontal line denotes the $P < 0.01$ genome-wide significance threshold. lod, log likelihood ratio. Error bars in **b** and **c**, s.e.m.

been shown to affect leaving, as well as locomotion speed on food, a behaviour that partially correlates with leaving rate[12,15] (Supplementary Fig. 2). Examining the *npr-1* genotype in the RIAILs revealed a strong but asymmetric correlation with leaving rates (Fig. 1c). Every strain with the N2 allele of *npr-1* had low leaving rates (≤1 event every 20 min), but strains with the HW allele of *npr-1* could have either low or high leaving rates (Fig. 1c). The asymmetric distribution is consistent with a role for *npr-1* in leaving behaviour, but indicates that *npr-1* has epistatic interactions with other loci segregating in the RIAILs.

The involvement of *npr-1* in leaving behaviour was confirmed by analysing near-isogenic lines (NILs) containing the N2 and HW *npr-1* alleles in the reciprocal strain background, and by examining *npr-1* null mutants (Supplementary Fig. 3). Specific transgenic expression of the N2 *npr-1* allele in its essential site of action, the RMG motor neurons[20], sharply reduced the leaving rate of HW animals (Supplementary Fig. 3). Thus *npr-1* is a regulator of HW leaving rates, but not the only contributing gene.

tyra-3 affects leaving behaviour

Studies in yeast, flies, mice, and plants have shown that individual QTLs often resolve into several genes that contribute to phenotypic variance[25–28]. Similarly, fine-mapping of the ~1 megabase (Mb) QTL that contained *npr-1* suggested the existence of multiple loci that affected leaving rates. A NIL with <150 kilobases (kb) of N2 DNA spanning the *npr-1* locus introgressed into HW had N2-like leaving rates (*leav-1* QTL, Fig. 2a and Supplementary Fig. 3). A second NIL with 700 kb of N2 DNA that did not cover *npr-1* introgressed into HW also had a low leaving rate, with about half the leaving rate of HW (*leav-2* QTL, Fig. 2a). These results suggest the existence of a second X-linked locus that affects leaving rates, which we called *leav-2*. The *leav-2* region did not affect leaving in the N2 genetic background (Fig. 2a), so all subsequent experiments were conducted in the HW background.

A 100 kb minimal region for *leav-2* was identified by analysing the breakpoints of individual RIAILs (Supplementary Fig. 4 and Supplementary Methods). We characterized the genetic properties of *leav-2* by crossing the *leav-2* NIL strain with HW. The heterozygous F_1 progeny had leaving rates similar to the *leav-2* NIL (Fig. 2a), indicating that the N2 *leav-2* locus was dominant to HW and suggesting that N2 transgenes covering the relevant gene should reduce the leaving rate of HW animals. Therefore, overlapping N2 genomic DNA fragments from the 100 kb minimal *leav-2* region were introduced into HW animals by microinjection (Fig. 2b and Supplementary Fig. 5). A single gene in this region reduced leaving rates: *tyra-3*, which encodes a G-protein-coupled receptor for the invertebrate noradrenaline-like neurotransmitters tyramine and octopamine[29]. Tyramine and octopamine receptors are related to vertebrate adrenergic receptors, and are thought to carry out analogous functions. *tyra-3* genomic fragments from the N2 strain were more active than *tyra-3* fragments from the HW strain injected at the same concentration, consistent with the possibility that *tyra-3* is a polymorphic gene that differs between N2 and HW (Fig. 2b).

If *leav-2* corresponds to *tyra-3*, a *tyra-3* mutation should eliminate its activity[30]. To test this prediction genetically, a null allele of *tyra-3* in an N2 background was introgressed into a HW background. The N2 region in the resulting NIL covered from 4.9 to 5.4 Mb of the X chromosome, the inferred position of *leav-2*. The *tyra-3(ok325)* null NIL had high (HW-like) leaving rates, suggesting that N2 *leav-2* activity was not present in the strain (Fig. 2a). Heterozygotes between HW and the near-isogenic *tyra-3(ok325)* null strain also had high leaving rates (Fig. 2a). These results are as expected if the active locus in *leav-2* is *tyra-3*; however, other genes within the introgressed regions could also contribute to the different leaving rates.

To strengthen the connection between *tyra-3* and *leav-2*, RNA interference against *tyra-3* was performed in the *leav-2* NIL that has low leaving rates due to the presence of the N2 QTL. Knockdown of *tyra-3* increased the leaving rate of the *leav-2* NIL to levels observed in HW animals, the result predicted if the *tyra-3* locus from N2 reduces leaving (Fig. 2c). Comparable experiments in a pure HW strain had minimal effects, as expected if *tyra-3* activity in HW is already low.

Further confirmation that the HW allele of *tyra-3* has reduced biological activity was provided by examining the one phenotype previously associated with *tyra-3*, avoidance of dilute octanol[29]. *tyra-3* null mutants avoid octanol more strongly than wild-type N2; the NIL strain with the HW *tyra-3* allele had a similar enhanced octanol response, suggesting that the HW *tyra-3* allele has reduced *tyra-3* function (Supplementary Fig. 6).

Non-coding changes affect *tyra-3* activity

The differential activity of N2 and HW genomic *tyra-3* fragments in the leaving assay suggested that N2 and HW alleles are functionally distinct (Fig. 2b). To identify polymorphisms between N2 and HW alleles of *tyra-3*, we sequenced ~19 kb surrounding the *tyra-3* locus in HW. There were 34 differences between HW and the N2 consensus genomic sequence (Fig. 3a): 33 non-coding changes and a single coding difference that changed a glutamic acid in the *tyra-3b* isoform to glycine.

Sequences that contribute to the differential activity of N2 and HW *tyra-3* alleles were localized further using transgenic assays. We fused

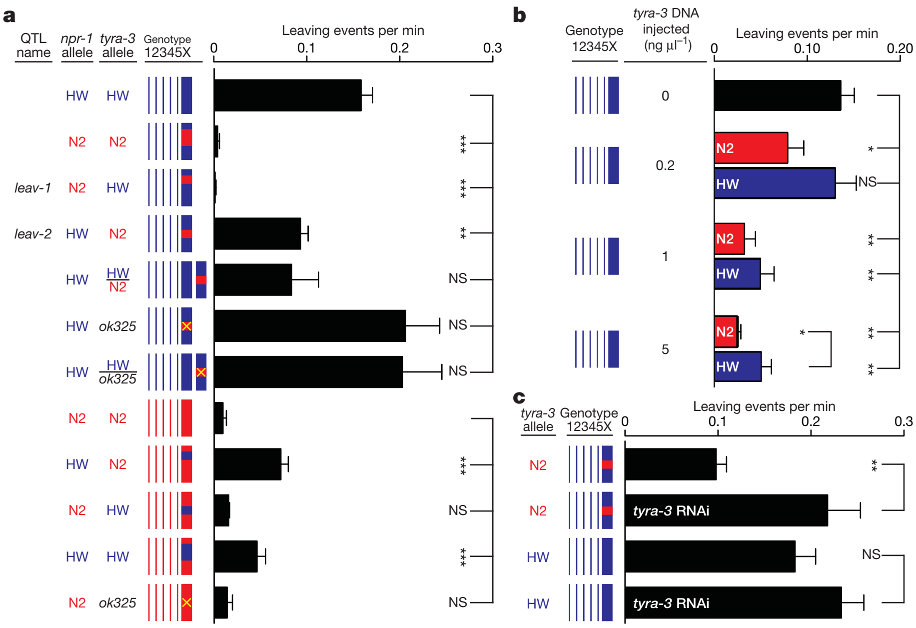


Figure 2 | N2 and HW *tyra-3* alleles differentially affect leaving rates. **a**, Dissection of the QTL on X into two loci: *leav-1* (4.70–4.78 Mb) and *leav-2* (4.78–5.75 Mb). 'Genotype' shows chromosomes; thick line is X chromosome. Blue denotes HW DNA, red denotes N2 DNA, and yellow denotes the *tyra-3(ok325)* null mutant. In heterozygous strains, both X chromosomes are shown. **b**, *tyra-3* genomic fragments (Fig. 3a) reduce HW leaving rates. Blue, HW transgenes; red, N2 transgenes. Two-way analysis of variance (ANOVA) showed significant effects of both transgene concentration and DNA strain of origin. **c**, Effect of *tyra-3* RNAi. Error bars, s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by *t*-test or ANOVA with Dunnett test. NS, not significant.

N2 and HW *tyra-3b* complementary DNAs to 4.9 kb of non-coding N2 or HW sequence upstream of the *tyra-3b* start site and introduced each of the four resulting clones into the HW strain. *tyra-3* transgenes with the N2 non-coding sequence were significantly more potent than comparable transgenes with the HW sequence, regardless of whether they preceded N2 or HW *tyra-3* cDNAs (Fig. 3b), excluding the coding polymorphism and localizing a functional difference between N2 and HW *tyra-3* genes to a 4.9 kb region that harbours 5 non-coding SNPs, 1 single nucleotide insertion, and a 184 bp deletion in HW. To narrow the relevant change down further, the 184 bp deletion was engineered into the N2 *tyra-3* genomic fragment; this clone was significantly less potent in the leaving assay than the full N2 genomic fragment (Supplementary Fig. 7). These results indicate that the 184 bp deletion represents at least part of the functional difference between N2 and HW *tyra-3* alleles.

Sequence variation in *tyra-3* non-coding regions could affect the level or location of *tyra-3* expression. Quantitative reverse transcription–PCR analysis (RT–PCR) of *tyra-3* mRNA levels in mixed-stage animals

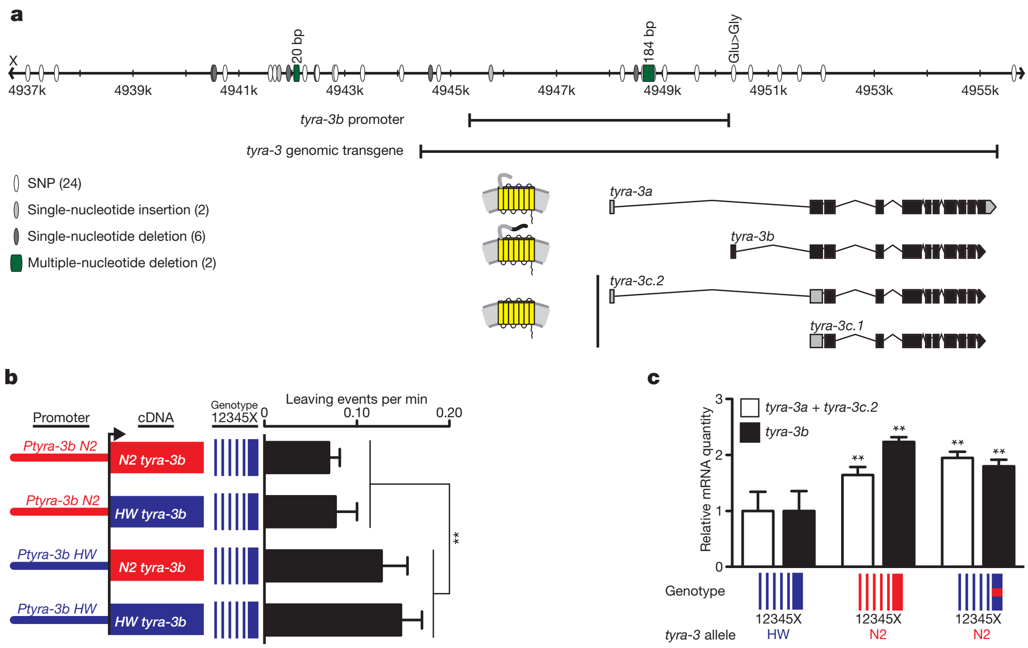


Figure 3 | Non-coding changes in *tyra-3* affect its activity and expression level. **a**, HW polymorphisms in the *tyra-3* locus relative to N2. *tyra-3* encodes three predicted G-protein-coupled receptors, shown as diagrams to the left of the gene structures for *tyra-3a*, *tyra-3b* and *tyra-3c*. The genomic region examined in Fig. 2b and the 4.9 kb promoter used in **b** and Fig. 4a are indicated. **b**, Leaving rates of transgenic HW animals with *tyra-3b* promoters (*Ptyra-3b*) fused to *tyra-3b* cDNAs. Error bars, s.e.m. **$P < 0.01$ by two-way ANOVA; no statistical interaction between the promoter and the cDNA. **c**, Relative amounts of *tyra-3* isoform mRNAs (**a**) in HW, N2 and *leav-2* strains (Fig. 2a). Error bars, s.d. **$P < 0.01$ compared to HW, ANOVA with Dunnett test.

indicated that N2 expressed approximately twice as much *tyra-3* mRNA as HW, consistent with increased *tyra-3* activity in the N2 strain (Fig. 3c). The *leav-2* NIL with N2 *tyra-3* introgressed into HW also had high *tyra-3* messenger RNA levels, suggesting that *cis*-acting changes affect *tyra-3* expression (Fig. 3c).

As both N2 and HW were cultivated in the laboratory for many years before permanent cultures were frozen, we wished to exclude the possibility that the *tyra-3* polymorphisms were laboratory-derived[19]. Therefore, 19 kb of the *tyra-3* locus was sequenced in all wild strains tested for leaving behaviour in Fig. 1, including three strains that were frozen immediately after their isolation. Each strain represents a different *C. elegans* haplotype group[24]. Both N2-like and HW-like *tyra-3* sequences were represented in the wild-caught strains, confirming the wild ancestry of both alleles (Supplementary Table 1 and Supplementary Methods). Notably, the *tyra-3* locus of MY1 was identical to N2 and, correspondingly, the leaving rate of MY1 was similar to that of N2.

tyra-3 acts in sensory neurons

The identification of *tyra-3* provided an opportunity to characterize the neuronal basis of the decision to leave or remain on a food patch. The biological activity of a transgene with 4.9 kb upstream of the *tyra-3b* start site fused to a *tyra-3* cDNA (Fig. 3b) implied that it was expressed in cells that regulate leaving behaviour. When this 4.9 kb region was fused to green fluorescent protein (GFP), it drove reliable expression in ASK, ADL, AIM, AUA, BAG, CEP, OLQ and SDQL neurons, in other unidentified neurons in the ventral ganglion and the tail, occasionally in AFD and AWC neurons, and in two non-neuronal cell types, the spermatheca and the distal tip cell (Fig. 4a and data not shown). The same set of cells was observed with reporter genes bearing either N2 or HW *tyra-3* upstream regions, and in both N2 and HW genetic backgrounds. Together with the quantitative RT–PCR data (Fig. 3c), these results suggest that different *tyra-3* expression levels, not different sites of expression, distinguish N2 and HW alleles.

The neurons whose activity is regulated by *tyra-3* were localized further by expressing *tyra-3* cDNAs from cell-type-specific promoters. *tyra-3* expression in ASK or BAG sensory neurons significantly reduced leaving, but expression in the CEP or ADL sensory neurons did not (Fig. 4b). The ASK neurons sense attractive food-derived amino acids[31] and regulate search behaviours after animals are removed from food[32,33]. The BAG neurons sense CO_2 and O_2, two cues associated with bacterial metabolism[34,35]. Lowering O_2 to levels that activate BAG reduced leaving rates (Supplementary Fig. 8).

To investigate whether the *tyra-3* non-coding polymorphism affects expression in relevant neurons, single-copy N2 or HW *tyra-3b* promoters driving GFP were inserted into a single, defined chromosomal location using the MosSCI technique[36]. GFP levels in ASK neurons were significantly higher for transgenes containing the N2 promoter compared to those containing the HW promoter (Fig. 4c). These results suggest that the N2 *tyra-3* locus is associated with higher *tyra-3* expression in ASK, as well as higher *tyra-3* mRNA expression at a whole-animal level; expression in BAG was not examined.

The behavioural functions of ASK and BAG, and the effect of *tyra-3* on those functions, were assessed by killing the neurons in different genetic backgrounds. Killing the ASK neurons reduced the leaving rate of HW animals, indicating that ASK can promote leaving (Fig. 4d). The ablation resembled the effect of the ASK::*tyra-3* transgene, suggesting that *tyra-3* reduces ASK activity. In agreement with this idea, killing the ASK neurons in a strain with the N2 high-activity *tyra-3* allele did not reduce their leaving rates further. The effect of *tyra-3* on ASK was selective for this assay; *tyra-3* did not reduce lysine chemotaxis, a second ASK-dependent behaviour (Supplementary Fig. 9).

Killing the BAG neurons increased leaving rates in the strain with the N2 *tyra-3* allele, demonstrating that BAG neurons prevent leaving (Fig. 4d). However, killing BAG had no effect in the strain with the HW *tyra-3* allele, suggesting that BAG activity is already low in this strain under the assay conditions. The ablation and genetic results suggest that the N2 *tyra-3* allele decreases ASK activity and increases BAG activity, two changes that act together to prevent leaving (Supplementary Fig. 1).

Gene–gene–environment interactions

Like most natural behaviours, the decision to leave a food patch is regulated by multiple genes and the environment; it responds to genetic variation in *tyra-3*, *npr-1*, and additional genes on the autosomes

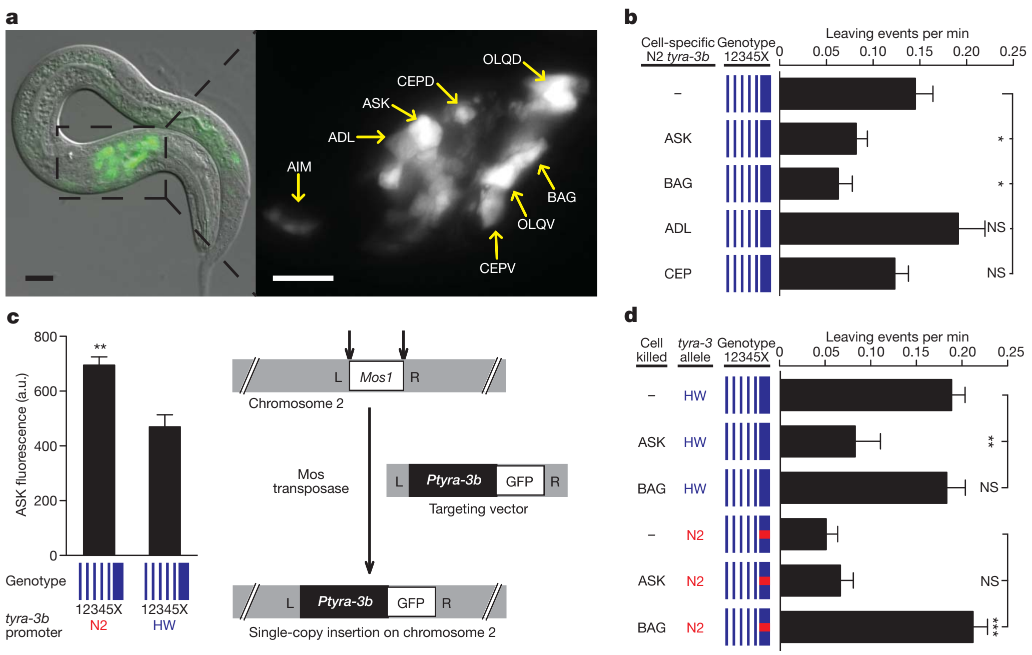


Figure 4 | *tyra-3* acts in ASK and BAG sensory neurons. **a**, Expression of 4.9 kb N2 *tyra-3b* promoter::GFP fusion (Fig. 3a) in HW animal at low (left) and high (right) magnification. HW *tyra-3b* promoter::GFP is expressed in the same cells, which are labelled in the right panel. Posterior signal is gut autofluorescence. Scale bar, 20 μm. **b**, Leaving rates of HW strains expressing *tyra-3b* in specific cells. **c**, Left, GFP fluorescence intensity in ASK of HW animals with a MosSCI insertion of N2 or HW 4.9 kb *tyra-3b* promoter::GFP. Right, schematic of MosSCI technique[36]. **d**, Leaving rates after killing ASK or BAG in HW and *leav-2* strains (Fig. 2a). Error bars, s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by *t*-test or ANOVA with Dunnett test.

(Fig. 1) as well as food quality and quantity[7,8]. Our results suggested that the N2 *npr-1* allele was epistatic to *tyra-3*; animals with the N2 *npr-1* allele had low leaving rates regardless of the *tyra-3* genotype (Fig. 2a). However, N2 *npr-1* reduced the leaving rate to almost zero, making it difficult to detect any further reduction. To make the assay more powerful, leaving was assayed on bacterial lawns of different densities. Leaving rates of all genotypes increased on thinner lawns and decreased on thicker lawns (Supplementary Fig. 10), but the thickness of the lawn changed the genetic interaction between *tyra-3* and *npr-1*. In the standard leaving assay, *tyra-3* polymorphisms had different effects only in the presence of the HW *npr-1* allele; on a thinner lawn, only in the presence of the N2 *npr-1* allele (Supplementary Fig. 10). Thus the epistatic relationship between *npr-1* and *tyra-3* is defined by the specific environment, not by an intrinsic regulatory relationship between the genes.

Discussion

Our results show that natural variation in *tyra-3* affects patch leaving, a behaviour representative of the exploration–exploitation decision. *tyra-3* encodes a G-protein-coupled receptor activated by the invertebrate transmitters tyramine and octopamine[29], which are structurally related to vertebrate adrenaline and noradrenaline. Catecholamines are known to regulate arousal systems that affect many behaviours and behavioural decisions. In *C. elegans,* octopamine drives sensory, molecular and behavioural responses to starvation, and tyramine affects specific aspects of locomotion[37–40]. In insects, octopamine affects locomotory activity, arousal and aggression[41–43]. Mammalian noradrenaline is generally implicated in arousal behaviours, and noradrenaline release from the locus coeruleus is associated with switching between different tasks, a cognitive function with analogies to the exploration–exploitation decision[6].

Relatively few natural behavioural variations have been mapped to the single-gene level in any animal, and it is interesting that several of these variations affect G-protein-coupled receptor signalling systems[44,45]. We speculate that these receptor pathways may serve as common substrates of behavioural variation. All animal genomes encode many G-protein-coupled receptors with different expression patterns. These receptors may provide a reservoir for genetic changes, as alteration in an individual receptor could cause relatively discrete effects without disrupting the entire system.

QTL mapping in rodents and in *Drosophila* indicates that most behavioural traits are polygenic, with widespread epistatic effects[1,46]. In agreement with this conclusion, our analysis suggests the existence of epistatic interactions between *tyra-3, npr-1* and at least one additional locus. Importantly, the non-additive interactions between *tyra-3* and *npr-1* are not stable, but vary based on the genetic background and the environment, similar to what has been found with yeast sporulation QTLs[47].

By integrating genetic studies of *C. elegans* foraging with neuronal analysis, we can provide a first-level description of underlying mechanisms. The sensory neurons that express *tyra-3* detect food-related cues; we suggest that they integrate these external cues with internal arousal states detected by *tyra-3*, and that different *tyra-3* alleles confer differential sensitivity to these arousal states (Supplementary Fig. 1). Thus variation in *tyra-3* lies at the intersection of many forms of behavioural flexibility: rapid responses to environmental cues, short-term modulation by internal state fluctuations, and long-term genetic changes that lead to adaptive changes in innate behaviours.

METHODS SUMMARY

Standardized leaving assays were conducted by videotaping seven adult hermaphrodite animals for 30 min on a bacterial lawn that was low-density compared to their growth lawn. The number of leaving events was recorded manually by examining the video recordings, and further behavioural analysis was conducted with a Matlab code adapted from the Parallel Worm Tracker[48]. A leaving event was defined as an episode in which the whole body of an animal left the bacterial lawn and the animal did not reverse immediately to return to the lawn. The leaving rate was calculated as the number of leaving events per worm minute spent inside the bacterial lawn. Experiments on each strain were repeated at least three times.

QTL analysis was performed on the mean leaving rates of N2-HW recombinant inbred advanced intercross lines (RIAILs) by non-parametric interval mapping in R/qtl[24,49]. Significance levels were estimated from 10,000 permutations of the data. Near-isogenic lines (NILs) were created by backcrossing a chromosomal region or allele into the desired genetic background at least 9 times. Extrachromosomal transgenes were made by injection of DNA clones into the gonads of young adult hermaphrodites together with a fluorescent coinjection marker[50]. To control for variation between transgenes, two to five independent lines from each injection were characterized. Single-copy insertion of transgenes was performed using the direct MosSCI transposition technique, targeting the *ttTi5605* Mos allele on chromosome II[36].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 March 2010; accepted 14 January 2011.
Published online 16 March 2011.

1. Flint, J. & Mackay, T. F. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* **19,** 723–733 (2009).
2. Stephens, D. W., Brown, J. S. & Ydenberg, R. C. *Foraging: Behavior and Ecology* (Univ. Chicago Press, 2007).
3. Charnov, E. L. Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9,** 129–136 (1976).
4. March, J. G. Exploration and exploitation in organizational learning. *Organ. Sci.* **2,** 71–87 (1991).
5. Barrett, H. C. & Fiddick, L. Evolution and risky decisions. *Trends Cogn. Sci.* **4,** 251–252 (2000).
6. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28,** 403–450 (2005).
7. Harvey, S. C. Non-dauer larval dispersal in *Caenorhabditis elegans. J. Exp. Zool. B* **312,** 224–230 (2009).
8. Shtonda, B. B. & Avery, L. Dietary choice behavior in *Caenorhabditis elegans. J. Exp. Biol.* **209,** 89–102 (2006).
9. Pujol, N. *et al.* A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans. Curr. Biol.* **11,** 809–821 (2001).
10. Pradel, E. *et al.* Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans. Proc. Natl Acad. Sci. USA* **104,** 2295–2300 (2007).
11. Lipton, J., Kleemann, G., Ghosh, R., Lints, R. & Emmons, S. W. Mate searching in *Caenorhabditis elegans*: a genetic model for sex drive in a simple invertebrate. *J. Neurosci.* **24,** 7427–7434 (2004).
12. Gloria-Soria, A. & Azevedo, R. B. *npr-1* Regulates foraging and dispersal strategies in *Caenorhabditis elegans. Curr. Biol.* **18,** 1694–1699 (2008).
13. Styer, K. L. *et al.* Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* **322,** 460–464 (2008).
14. Reddy, K. C., Andersen, E. C., Kruglyak, L. & Kim, D. H. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans. Science* **323,** 382–384 (2009).
15. de Bono, M. & Bargmann, C. I. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans. Cell* **94,** 679–689 (1998).
16. Gray, J. M. *et al.* Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature* **430,** 317–322 (2004).
17. Rogers, C., Persson, A., Cheung, B. & de Bono, M. Behavioral motifs and neural pathways coordinating O_2 responses and aggregation in *C. elegans. Curr. Biol.* **16,** 649–659 (2006).
18. Bretscher, A. J., Busch, K. E. & de Bono, M. A carbon dioxide avoidance behavior is integrated with responses to ambient oxygen and food in *Caenorhabditis elegans. Proc. Natl Acad. Sci. USA* **105,** 8044–8049 (2008).
19. McGrath, P. T. *et al.* Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61,** 692–699 (2009).
20. Macosko, E. Z. *et al.* A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans. Nature* **458,** 1171–1175 (2009).
21. Osborne, K. A. *et al.* Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila. Science* **277,** 834–836 (1997).
22. Reaume, C. J. & Sokolowski, M. B. cGMP-dependent protein kinase as a modifier of behaviour. *Handb. Exp. Pharmacol.* **191,** 423–443 (2009).
23. Hong, R. L., Witte, H. & Sommer, R. J. Natural variation in *Pristionchus pacificus* insect pheromone attraction involves the protein kinase EGL-4. *Proc. Natl Acad. Sci. USA* **105,** 7779–7784 (2008).
24. Rockman, M. V. & Kruglyak, L. Recombinational landscape and population genomics of *Caenorhabditis elegans. PLoS Genet.* **5,** e1000419 (2009).
25. Steinmetz, L. M. *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416,** 326–330 (2002).
26. Edwards, A. C. & Mackay, T. F. Quantitative trait loci for aggressive behavior in *Drosophila melanogaster. Genetics* **182,** 889–897 (2009).

27. Legare, M. E., Bartlett, F. S. II & Frankel, W. N. A major effect QTL determined by multiple genes in epileptic EL mice. *Genome Res.* **10,** 42–48 (2000).
28. Thomson, M. J., Edwards, J. D., Septiningsih, E. M., Harrington, S. E. & McCouch, S. R. Substitution mapping of *dth1.1*, a flowering-time quantitative trait locus (QTL) associated with transgressive variation in rice, reveals multiple sub-QTL. *Genetics* **172,** 2501–2514 (2006).
29. Wragg, R. T. *et al.* Tyramine and octopamine independently inhibit serotonin-stimulated aversive behaviors in *Caenorhabditis elegans* through two novel amine receptors. *J. Neurosci.* **27,** 13402–13412 (2007).
30. Mackay, T. F. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2,** 11–20 (2001).
31. Bargmann, C. I. & Horvitz, H. R. Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans*. *Neuron* **7,** 729–742 (1991).
32. Gray, J. M., Hill, J. J. & Bargmann, C. I. A circuit for navigation in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **102,** 3184–3191 (2005).
33. Wakabayashi, T., Kitagawa, I. & Shingai, R. Neurons regulating the duration of forward locomotion in *Caenorhabditis elegans*. *Neurosci. Res.* **50,** 103–111 (2004).
34. Hallem, E. A. & Sternberg, P. W. Acute carbon dioxide avoidance in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **105,** 8038–8043 (2008).
35. Zimmer, M. *et al.* Neurons detect increases and decreases in oxygen levels using distinct guanylate cyclases. *Neuron* **61,** 865–879 (2009).
36. Frøkjær-Jensen, C. *et al.* Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nature Genet.* **40,** 1375–1383 (2008).
37. Suo, S., Kimura, Y. & Van Tol, H. H. Starvation induces cAMP response element-binding protein-dependent gene expression through octopamine-Gq signaling in *Caenorhabditis elegans*. *J. Neurosci.* **26,** 10082–10090 (2006).
38. Alkema, M. J., Hunter-Ensor, M., Ringstad, N. & Horvitz, H. R. Tyramine functions independently of octopamine in the *Caenorhabditis elegans* nervous system. *Neuron* **46,** 247–260 (2005).
39. Greer, E. R., Perez, C. L., Van Gilst, M. R., Lee, B. H. & Ashrafi, K. Neural and molecular dissection of a *C. elegans* sensory circuit that regulates fat and feeding. *Cell Metab.* **8,** 118–131 (2008).
40. Pirri, J. K., McPherson, A. D., Donnelly, J. L., Francis, M. M. & Alkema, M. J. A tyramine-gated chloride channel coordinates distinct motor programs of a *Caenorhabditis elegans* escape response. *Neuron* **62,** 526–538 (2009).
41. Crocker, A. & Sehgal, A. Octopamine regulates sleep in *Drosophila* through protein kinase A-dependent mechanisms. *J. Neurosci.* **28,** 9377–9385 (2008).
42. Roeder, T. Tyramine and octopamine: ruling behavior and metabolism. *Annu. Rev. Entomol.* **50,** 447–477 (2005).
43. Hoyer, S. C. *et al.* Octopamine in male aggression of *Drosophila*. *Curr. Biol.* **18,** 159–167 (2008).
44. Yalcin, B. *et al.* Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nature Genet.* **36,** 1197–1202 (2004).
45. Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R. & Insel, T. R. Increased affiliative response to vasopressin in mice expressing the V_{1a} receptor from a monogamous vole. *Nature* **400,** 766–768 (1999).
46. Mackay, T. F., Stone, E. A. & Ayroles, J. F. The genetics of quantitative traits: challenges and prospects. *Nature Rev. Genet.* **10,** 565–577 (2009).
47. Gerke, J., Lorenz, K., Ramnarine, S. & Cohen, B. Gene-environment interactions at nucleotide resolution. *PLoS Genet.* **6,** e1001144 (2010).
48. Ramot, D., Johnson, B. E., Berry, T. L. Jr, Carnell, L. & Goodman, M. B. The Parallel Worm Tracker: a platform for measuring average speed and drug-induced paralysis in nematodes. *PLoS ONE* **3,** e2208 (2008).
49. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).
50. Mello, C. & Fire, A. DNA transformation. *Methods Cell Biol.* **48,** 451–482 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Shingai for strains, and P. McGrath and members of the C.I.B. laboratory for discussions. A.B. was supported by the Secretaría de Educación Pública of Mexico and by The Rockefeller University. C.I.B. and L.K. are Investigators of the Howard Hughes Medical Institute. This work was supported by HHMI and by NIH grant GM089972.

Author Contributions A.B. and C.I.B. designed experiments, A.B. conducted experiments, M.V.R. constructed strains for QTL mapping, M.T. developed tracking methods, A.B., M.V.R., L.K. and C.I.B. analysed and interpreted results, and A.B. and C.I.B. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.I.B. (cori@rockefeller.edu).

METHODS

Analysis of behaviour in the leaving assay. Six-centimetre NGM agar plates were seeded with 70 μl (conditioning plate) or with 10 μl (assay plate) of a fresh overnight culture of *Escherichia coli* HB101 diluted in LB to $A_{600nm} = 2.0$. 90 min after seeding the plates, ten young adult hermaphrodites were picked onto the conditioning plate. 30 min after being placed on the conditioning plates, seven of the animals were transferred onto the lawn of the assay plate. The 30 min leaving assay began 1 h after placing the seven animals on the assay plate. The number of leaving events was recorded manually by examining the video recordings, and further behavioural analysis was conducted with a Matlab code adapted from the Parallel Worm Tracker[48]. A leaving event was defined as an episode in which the whole body of an animal left the bacterial lawn and the animal did not reverse immediately to return to the lawn. The leaving rate was calculated as the number of leaving events per worm minute spent inside the bacterial lawn. Experiments on each strain were repeated at least three times.

Quantitative trait locus analysis. The N2-HW recombinant inbred advanced intercross lines (RIAILs) used in this study represent the terminal generation of a 20-generation pedigree founded by reciprocal crosses between N2 and HW. The lines were constructed through 10 generations of intercrossing followed by 10 generations of selfing[24]. They have been genotyped at 1,454 nuclear and one mitochondrial markers and have a 5.3-fold expansion of the F_2 genetic map[24]. QTL analysis was performed on the mean leaving rates of N2-HW RIAILs by non-parametric interval mapping in R/qtl[49]. Significance levels were estimated from 10,000 permutations of the data.

Quantitative RT–PCR. Total RNA from mixed stage worms was isolated with Trizol. 1.5 μg of RNA and oligo-dT were used for reverse transcription using SuperScript III First-Strand Synthesis (Invitrogen) according to the manufacturer's instructions. Real-time PCR was performed with Fast SYBR Green Master Mix (Applied Biosystems) on a 7900HT Real-Time PCR System (Applied Biosystems). *act-3* was used as the calibrator for relative quantitation. 5′ primers corresponded to upstream exons that distinguished *tyra-3* isoforms, and 3′ primers corresponded to shared exon sequence. Primers used were: *tyra-3a & tyra-3c.2_F,* ccacttgcaaatagcagcag; *tyra-3b_F,* ggctatttggtggtggtttg; *tyra-3a & tyra-3b_R,* tccttctggcgtcgaaatac; *act-3_F,* tcacgatcatgagaccattcaaa; *act-3_R,* gcaaattgtagtggggtcttcttatg.

***tyra-3* expression pattern.** The N2 and HW 4.9 kb *tyra-3b* promoters were amplified using primers: tcaacctaaccactaactaaggg and cGatgaagcaagatgtcaggt, which overlaps the coding region by 4 bp. The ATG start codon is mutated to ATC (mutation is uppercase in primer). These promoters were individually fused by PCR to a fragment containing GFP followed by the *unc-54* 3′-UTR, as described[51]. These PCR products were injected individually into both HW and N2 animals at 20 ng μl^{-1}. Cells expressing GFP were identified by Nomarski microscopy in both L1 and adult hermaphrodites. The identification of some cells was aided by injecting *Ptyra-3b*::GFP-expressing animals with promoter-mCherry fusions with established expression patterns. In this manner, the AIM neurons were identified as *Ptyra-3b*::GFP-expressing cells based on their position and the absence of co-localization with *Pttx-3*::mCherry. The BAG neurons co-expressed *Ptyra-3b*::GFP and *Pflp-17*::mCherry. The CEP neurons co-expressed *Ptyra-3b*::GFP and *Pdat-1*::mCherry. The ASK neurons co-expressed *Ptyra-3b*::GFP and *Psra-9*::mCherry. The ADL neurons co-expressed *Ptyra-3b*::GFP and *Psri-51*::mCherry.

Extrachromosomal transgenes. Transgenes were made by injection of DNA clones into the gonads of young adult hermaphrodites together with a fluorescent coinjection marker[50]. To control for variation between transgenes, two to five independent lines from each injection were characterized.

Generation of MosSCI lines and quantitation of GFP fluorescence in ASK. Single-copy insertion of transgenes was performed using the direct MosSCI technique targeting the *ttTi5605* Mos allele on chromosome II, as described[36]. A schematic of the mechanism underlying MosSCI is shown in Fig. 4c.

The pCFJ151 targeting vector was modified by the introduction of an FseI restriction site into the multiple cloning site by site-directed mutagenesis using the primers gtaatacgactcacttaaggccggccctagagggtaccagagctcacc and ggtgagctctggt accctctagggccggccttaagtgagtcgtattac to make pAB1. An FseI-SpeI fragment from a pSM vector containing *N2-Ptyra-3b::N2-tyra-3b::SL2 GFP::unc-54 3′-UTR* or *HW-Ptyra-3b::N2-tyra-3b::SL2 GFP::unc-54 3′-UTR* was cloned into pAB1.

For each *tyra-3*-containing test plasmid, about 50 EG4322 animals were injected with a mixture of *tyra-3* plasmid, pGH8, pCFJ90, pCFJ104 and pJL43.1. After positive and negative selection and full sequencing of the insert, two inserted transgenes each of N2-*Ptyra-3b* and HW-*Ptyra-3b* were backcrossed to HW males seven times, selecting GFP-fluorescent hermaphrodites each generation. The transgene-containing chromosome was then homozygosed.

The strains containing the single-copy transgene in a HW background were injected with *Psra-9*::mCherry to identify ASK. Young adult hermaphrodites were examined on a Zeiss Imager Z.1 with a ×60 objective focused on ASK using mCherry to prevent bleaching of GFP signal. Fluorescence signals were acquired with fixed acquisition times (30–50 ms for mCherry, 100 ms for GFP). Background mean fluorescence intensity adjacent to ASK was subtracted from the ASK signal.

RNAi. RNA interference was performed essentially as described[52]. A fragment common to all *tyra-3* isoforms was amplified. The following primers were used, which include the T7 sequence (underlined): <u>taatacgactcactatagggaga</u>gaaaatggc agcaggacttt; <u>taatacgactcactatagggaga</u>atcctcgcagtctgtggagt. *in vitro* transcription was performed with RiboMAX kit (Promega). dsRNA was injected at 1.2 μg μl^{-1} into the gonads of adult hermaphrodites. Eggs laid 24 and 48 h after injection were used for the behavioural assays.

Octanol avoidance assay. Avoidance assays were conducted essentially as described[53]. In brief, ~20 three-day-old animals were picked off of their growth plates food into a transfer plate without bacteria where they were allowed to crawl and rid themselves of bacteria. Animals were then transferred onto an NGM plate without food. After 40 min, a microcapillary with 30% octanol (v/v diluted fresh every day in ethanol) was presented in front of the animal's nose. The time to reverse was recorded. If animals did not reverse within 20 s, the assay was stopped. Animals were presented with odour 1–3 times per experiment, with at least 3 min of rest interval. We replicated published results demonstrating that *tyra-3* null mutants had more rapid responses than N2 in the presence of exogenous serotonin and tyramine[29] but also observed more rapid responses in the absence of exogenous neuromodulators, as shown in Supplementary Fig. 6.

Cell ablations. For leaving behaviour assays, ASK was ablated with a laser microbeam as described[54]. BAG was killed using split human caspase 3 fragments[55] expressed from *flp-17* and *glb-5* promoters that overlapped only in BAG. For lysine chemotaxis assays, ASK was killed using a mouse caspase 1 gene expressed from the *sra-9* promoter[56]. The ASK strain was a gift from R. Shingai.

51. Hobert, O. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechnique* **32,** 728–730 (2002).
52. Ahringer, J. Reverse genetics. In *WormBook* (ed. The *C. elegans* Research Community) doi:10.1895/wormbook.1.47.1 (6 April 2006); available at ⟨http://www.wormbook.org⟩.
53. Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. A. & Bargmann, C. I. Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*. *Cell* **83,** 207–218 (1995).
54. Bargmann, C. I. & Avery, L. Laser killing of cells in *Caenorhabditis elegans*. *Methods Cell Biol.* **48,** 225–250 (1995).
55. Chelur, D. S. & Chalfie, M. Targeted cell killing by reconstituted caspases. *Proc. Natl Acad. Sci. USA* **104,** 2283–2288 (2007).
56. Kim, K. *et al.* Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **326,** 994–998 (2009).

doi:10.1038/nature09788

Caspase signalling controls microglia activation and neurotoxicity

Miguel A. Burguillos[1,2,3], Tomas Deierborg[3], Edel Kavanagh[1], Annette Persson[4], Nabil Hajji[1]†, Albert Garcia-Quintanilla[5], Josefina Cano[2], Patrik Brundin[3], Elisabet Englund[4], Jose L. Venero[2] & Bertrand Joseph[1]

Activation of microglia and inflammation-mediated neurotoxicity are suggested to play a decisive role in the pathogenesis of several neurodegenerative disorders. Activated microglia release pro-inflammatory factors that may be neurotoxic. Here we show that the orderly activation of caspase-8 and caspase-3/7, known executioners of apoptotic cell death, regulate microglia activation through a protein kinase C (PKC)-δ-dependent pathway. We find that stimulation of microglia with various inflammogens activates caspase-8 and caspase-3/7 in microglia without triggering cell death *in vitro* and *in vivo*. Knockdown or chemical inhibition of each of these caspases hindered microglia activation and consequently reduced neurotoxicity. We observe that these caspases are activated in microglia in the ventral mesencephalon of Parkinson's disease (PD) and the frontal cortex of individuals with Alzheimer's disease (AD). Taken together, we show that caspase-8 and caspase-3/7 are involved in regulating microglia activation. We conclude that inhibition of these caspases could be neuroprotective by targeting the microglia rather than the neurons themselves.

Numerous *in vivo* clinical imaging and neuropathology studies suggest that activated microglia, the resident immune cells of the central nervous system, play prominent roles in the pathogenesis of neurodegenerative disorders, including PD, multiple sclerosis and AD[1,2]. Microglia are necessary for normal brain function; however, uncontrolled and over-activated microglia can trigger neurotoxicity. They are a prominent source of pro-inflammatory factors and oxidative stress such as tumour-necrosis factor (TNF)-α, nitric oxide and interleukin (IL)-1β, which are neurotoxic[2,3].

Toll-like receptors (TLRs) are a family of pattern-recognition receptors in the innate immune system. Exogenous and endogenous TLR ligands activate microglia[1]. Intracerebral delivery of lipopolysaccharide (LPS), the major component of Gram-negative bacterial walls and a ligand for TLR4, leads *in vivo* to microglia activation and neuronal injury, and is used as model for brain inflammation[4,5]. Synergistic effects between interferon-γ (IFN-γ) and several TLR ligands (including TLR4) have been suggested, suggesting crosstalk between these pro-inflammatory receptor signalling pathways[6]. Furthermore, IFN-γ receptor-deficient mice are less susceptible to LPS-induced endotoxic shock than control mice[7]. Finally, TLR4 has been implicated in AD pathophysiology in several contexts. Thus, the upregulation of cytokines is TLR4 dependent in an AD mouse model[8]; certain TLR4 single nucleotide polymorphisms are associated with increased risk for AD[9]; the levels of TLR4 messenger RNA (mRNA) are upregulated in APP transgenic mice[10]; and increased TLR4 expression is associated with amyloid plaque deposition in AD brain tissue[10].

Caspases, a family of cysteinyl-aspartate-specific proteases, are executioners of apoptotic cell death and their activation is considered a commitment to cell death[11,12]. Certain caspases, for example caspase-1, also play a pivotal role in immune-mediated inflammation. In this situation, caspase activation is associated with the maturation of pro-inflammatory cytokines, such as IL-1β, IL-18, IL-33, and not with apoptosis[13]. Inhibition of caspase activation protects against neuronal loss in several animal models of brain diseases involving activated microglia, including hypoxic ischaemia/stroke, acute bacterial meningitis, brain trauma and 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-lesioned parkinsonism models[2,14–17]. Currently, it is unclear whether inhibition of caspase activation specifically in microglia contributes to the neuroprotective effects of caspase inhibitors. We have now discovered that microglial activation in cell and animal models of inflammation involves caspases and that inhibition of the cascade in microglia prevents neurodegeneration. Furthermore, we demonstrate that caspase activation occurs in microglia in the brains of individuals with PD and AD, and thereby we validate the observations we made in relevant cell and animal models.

Results

Caspase-3/7 control microglia activation

We stimulated BV2 cells with LPS to investigate the molecular pathways involved in microglia activation (Fig. 1). LPS treatment induced caspase-3 cleavage and D(OMe)E(OMe)VD(OMe)-ase (DEVD-ase) activity after only 4 h in BV2 microglia cells in a time- and dose-dependent manner (Fig. 1a and Supplementary Figs 2a, b, e, f and 5d, e). Increased DEVD-ase activity was also observed upon treatment with other pro-inflammogens such as lipoteichoic acid (LTA, TLR2 agonist), PamC3sk4 (synthetic lipopeptide TLR1/2 agonist) and interferon-gamma (IFN-γ) (Supplementary Fig. 3a). This activity primarily reflects caspase-3/7 activities. Both caspase 3 and 7 are known as major apoptosis executioners. Despite the increase in DEVD-ase activity, we did not observe major microglia cell death within the first 24 h after initiating LPS treatment (Fig. 1c and Supplementary Figs 2c and 4a–c, e, f) nor with LTA, PamC3sk4 and IFN-γ treatments (Supplementary Fig. 3b, c). Thus, the LPS-induced increase in caspase 3 and 7 activity did not result in major cell death and the little cell death occurring at 48 h was not prevented using the caspase-3/7 inhibitor Z-D(OMe)E(OMe)VD(OMe)-FMK

[1]Department of Oncology-Pathology, Cancer Centrum Karolinska, Karolinska Institutet, 171 76, Stockholm, Sweden. [2]Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Sevilla, and Instituto de Biomedicina de Sevilla, 41012 Sevilla, Spain. [3]Neuronal Survival Unit, Wallenberg Neuroscience Center, Department of Experimental Medical Science, 221 84 Lund, Sweden. [4]Department of Pathology, Division of Neuropathology, Lund University Hospital, 221 85 Lund, Sweden. [5]Servicio de Biología, Centro de Investigación, Tecnología e Innovación, Universidad de Sevilla (CITIUS), 41012 Sevilla, Spain. †Present address: Department of Experimental Medicine and Toxicology, Division of Investigative Science, Imperial College London, UK.

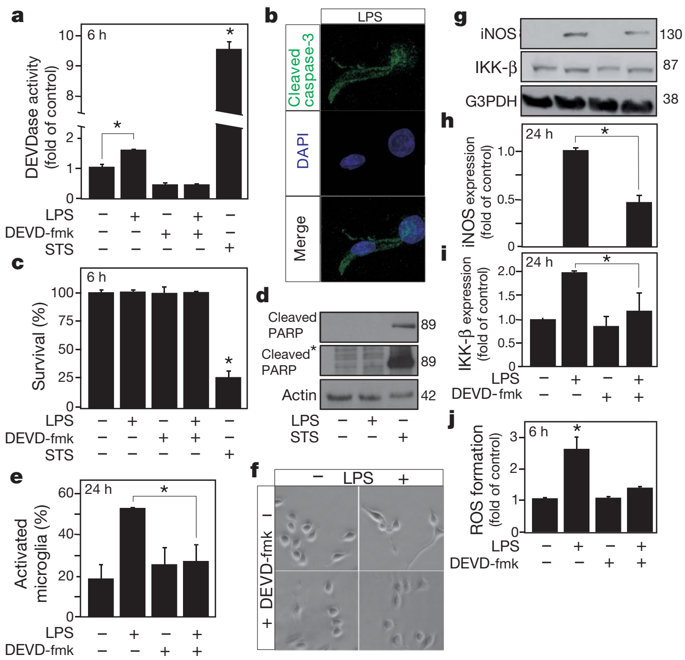


Figure 1 | LPS-induced DEVDase activity regulates microglia activation but not cell death. LPS treatment induces DEVDase activity (**a**) and processing of caspase-3 (**b**), which are not associated with cell death, as illustrated by cell survival quantification (**c**), and absence of PARP cleavage (**d**) in cultured BV2 microglia cells. Caspase-3/7 inhibition by DEVD-fmk (**a**) prevents LPS-induced morphological microglia activation (**e**, **f**) and induction of typical inflammation-related molecules like iNOS (**g**, **h**) and IKK-β (**g**, **i**) and ROS formation (**j**). STS is used as an apoptosis inducer. Data are expressed as mean ± s.e.m. ($n = 3$). $^{*}P < 0.05$.

(DEVD-fmk) (Supplementary Figs 2d and 4d, g). Compared with LPS treatment, exposure of BV2 microglia cells to a death stimulus such as staurosporine (STS) led to a significantly greater caspase-3 cleavage and induction of DEVD-ase activity (Fig. 1a and Supplementary Fig. 5d, e). After LPS treatment, we found cleaved caspase-3 to be located primarily close to the plasma membrane and not present in the nucleus (Fig. 1b, Supplementary Fig. 5a–d and Supplementary Movies 1 and 2). Furthermore, we did not observe cleavage of the caspase-3/7 nuclear substrate poly(ADP-ribose) polymerase (PARP-1) in response to LPS (Fig. 1d) even after long exposure of the membrane (asterisk in Fig. 1d). In addition, whereas exposure of BV2 cells to STS promoted Bid processing and loss of mitochondrial transmembrane potential, these two events associated with the mitochondrial cell death pathway were found to be unaffected upon LPS treatment (Supplementary Fig. 10a–c). When we inhibited DEVD-ase activity in BV2 cells by exposing them to the cell-permeable and irreversible caspase inhibitor DEVD-fmk, LPS treatment failed to activate the microglia. Thus, the cells did not exhibit morphological changes associated with microglia activation (Fig. 1e–f) and did not show features of activated microglia, such as IκB kinase complex β (IKK-β), inducible nitric oxide synthases (iNOS) and reactive oxygen species (ROS) formation (Fig. 1g–j). Co-treatment with DEVD-fmk also prevented the LTA, PamC3sk4 and IFN-γ-induced iNOS expression and ROS formation (Supplementary Fig. 6a–e).

Caspase-3 and -7 account for cellular DEVD-ase activity. We therefore decided to assess their respective roles in microglia activation by selectively knocking down endogenous caspase-3 (Supplementary Fig. 7a, b) or caspase-7 (Supplementary Fig. 7c, d) using a pool of small interfering RNAs (siRNAs). First, we confirmed that the silencing of these proteases effectively decreased DEVD-ase activity (Supplementary Fig. 7e). Then, when we transfected BV2 microglia cells with siRNA targeting specifically either one of the two caspases, LPS-treatment did not induce iNOS, IKK-β expression, ROS formation and production of certain cytokine production (IL-1β, TNF-α and murine keratinocyte chemoattractant) as effectively (Fig. 2a, b and Supplementary Fig. 7f–i). We found that silencing both caspases simultaneously reduced IKK-β expression even further (Supplementary Fig. 7f). This indicates that DEVD-ase activity per se regulates microglia activation (Fig. 2 and Supplementary Fig. 7). Similar responses to caspase-3 and -7 knockdowns were observed with LTA, PamC3sk4 and IFN-γ treatments (Supplementary Fig. 8a–i). Activation and nuclear translocation of nuclear factor κB (NF-κB) is a key step in LPS-induced microglia activation[18]. NF-κB is sequestered in the cytoplasm by the IκB family of inhibitory proteins that mask the nuclear localization signal of NF-κB. IKK-β can phosphorylate IκB, and thereby target it for degradation through the ubiquitin proteasome pathway. As a consequence, functional NF-κB molecules then become free to enter the nucleus. We detected less nuclear NF-κB p65 subunits in cells subjected to knockdown of caspase-3 or caspase-7 before LPS treatment (Fig. 2c, d), indicating that reduced caspase activation led to less nuclear translocation of NF-κB. Finally, using microglia cells co-cultured with dopaminergic neurons, we examined whether the inhibition of the IKK/NF-κB pathways by selective knockdown of caspase-3 or/and capase-7 was associated with loss of microglia neurotoxicity. In agreement with earlier studies[19], LPS treatment activated microglia and caused dopaminergic neurons to die. Importantly, we found that reducing the LPS-induced microglia activation by caspase knockdown meant that fewer co-cultured dopaminergic neurons died (Fig. 2e, f). Also, we checked the level of several cytokines in primary microglial cell cultures at 12 and 24 h, and observed a decrease of IL-1β and IL-5 with both DEVD-fmk and IETD-fmk, and of IL-2 and IFN-γ at 12 h when we used IETD-fmk (Fig. 2g).

Caspase-8 directs caspase-3/7 activation

We then examined how caspase-3/7 are activated in LPS-treated microglia. LPS treatment has been reported to promote caspase-1 activation[20]. Caspase 1 is a key component of the inflammasome required for the processing and maturation of pro-inflammatory cytokines. It plays a pivotal role during LPS-induced inflammation. Therefore caspase-1 null mice and mice expressing a dominant-negative mutant caspase-1 gene exhibit reduced LPS-induced inflammation[21,22]. Consequently, we examined if caspase-1 acts upstream of caspase-3/7 in LPS-induced microglia activation. We found that the caspase-1 inhibitor YVAD-fmk did not reduce LPS-induced activation of microglia, as assessed by DEVD-ase activity (Fig. 3a).

Caspase-8 is believed to be at the apex of the death receptor-mediated apoptosis pathway and can activate caspase-3/7 (refs 23–25). We found that LPS induced caspase-8 activity (IETD-ase) within 6 h of being added to BV2 microglia cultures (Fig. 3b, c). Consistent with this result, we observed that the caspase-8 inhibitor IETD-fmk or caspase-8 knockdown using siRNA prevented LPS-induced DEVD-ase activity (Fig. 3a, d). The IETD-fmk treatment or caspase-8 silencing only prevented the LPS-induced increase in DEVD-ase activity, whereas the caspase inhibitor DEVD-fmk reduced the DEVD-ase activity even further (Fig. 3a, c). Knockdown or chemical inhibition of caspase-8 were also associated with a reduction of iNOS expression and ROS formation upon treatment with all tested pro-inflammogens (that is, LPS, LTA, PamC3sk4 and IFN-γ) (Fig. 3e, f and Supplementary Figs 7i, 8a–i and 9a, b). Other known potential caspase-8 substrates, namely Bid, HDAC7 and RIP1, were found not to be processed upon LPS treatment (Supplementary Fig. 10 b–f). Taken together, this indicates that the LPS-induced increase in DEVD-ase activity is dependent on caspase-8 activation. Furthermore, we found that the caspase-8- and caspase-1-initiated pathways had additive effects in the regulation of LPS-induced iNOS expression (Supplementary Fig. 9c, d). This suggests that both of these caspase-regulated pathways contribute to LPS-induced inflammation. We also examined if the observed DEVD-ase activity and that of caspase 6, the remaining executioner caspase, were correlated. However, we were unable to

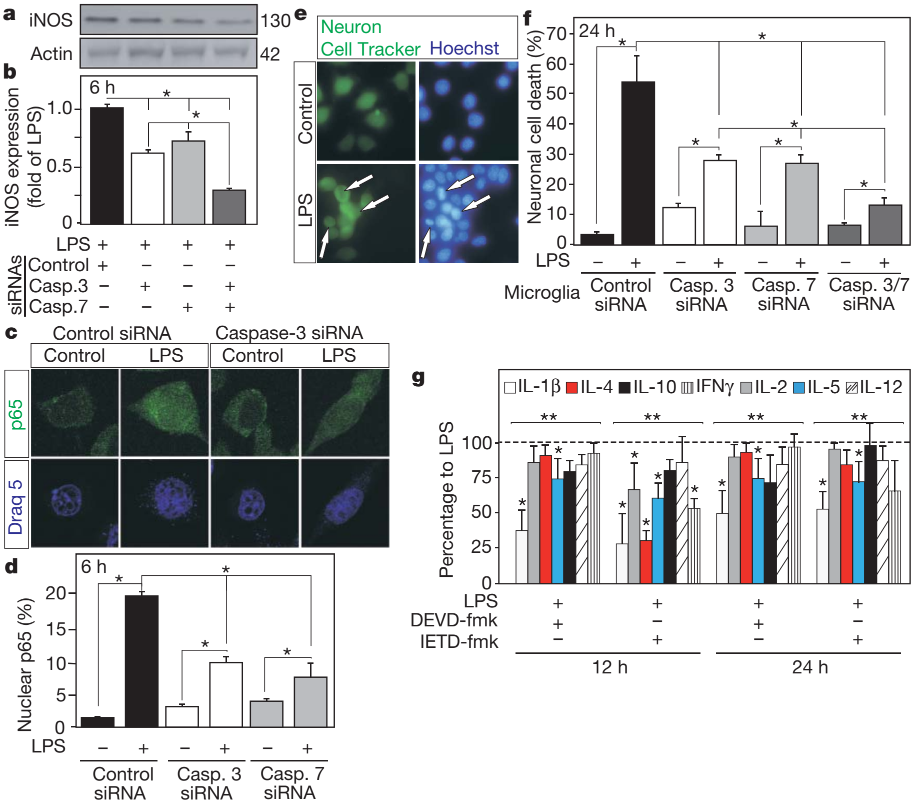


Figure 2 | Knockdown of caspase-3 or caspase-7 decreases microglia activation in response to LPS. siRNAs targeting caspase-3 and/or caspase-7 prevent LPS-induced increase of iNOS expression (**a**, **b**) in cultured BV2 cells. In addition, caspase-3 or caspase-7 silencing prevents LPS-induced activation of NF-κB as seen by reduction of p65 nuclear translocation (**c**, **d**). Under co-cultured conditions, LPS treatment induces neuronal cell death and caspase-3/7 knockdown in microglia cells prevents neuronal cell death (**e**, **f**). DEVD-fmk or IETD-fmk (caspase-8 inhibitor) treatment modulates cytokine expression at 12 and 24 h in primary microglia cell culture (**g**). Data are expressed as mean ± s.e.m. ($n = 3$) and ± s.d. ($n = 4$) in **g**. *$P < 0.05$. In **g**, ** denotes statistically significant difference between treatments ($P < 0.001$) concerning time and treatment.

detect any VEID-ase activity (related to caspase-6) in cells upon LPS treatment. Pretreatment with VEID-fmk, a caspase 6 inhibitor, did not affect LPS-induced DEVD-ase activity, indicating that this protease does not play an essential role in the activation process (Supplementary Fig. 11a, b).

The caspase signalling depends on TLR4

We then examined further the link between TLR4 ligation and activation of caspase-8. Selective knockdown of TLR4 (Supplementary Fig. 12a) was associated with a reduced caspase-8 activation and consequent caspase-3 activation, providing evidence for a direct activation of caspase-8 by TLR4 ligation (Supplementary Fig. 12b, c). Apoptosis after TLR2 activation has been reported to be associated to the formation of a myeloid differentiation factor 88 (MyD88)/Fas-associated

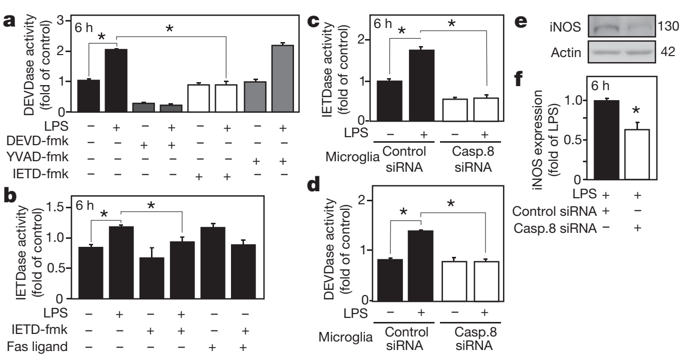


Figure 3 | Caspase-8 activity controls LPS-induced caspase-3/7 activation. Caspase-8 inhibition by IETD-fmk, but not caspase-1 inhibition by YVAD-fmk, prevents LPS-induced DEVDase activity (**a**). LPS treatment induces caspase-8 (IETDase) activity, which is inhibited by IETD-fmk (**b**) or specific siRNA targeting caspase-8 (**c**) in BV2 cells. Treatment with Fas ligand was used as positive control for IETDase activity induction (**b**). Caspase-8 knockdown using siRNA prevents LPS-induced DEVDase activity (**d**). siRNA knockdown of caspase-8 prevents LPS-induced expression of iNOS (**e**, **f**). Data are expressed as mean ± s.e.m. ($n = 3$). *$P < 0.05$.

death domain protein (FADD)/Caspase-8 complex[26]. To assess the potential role of this complex in TLR4 ligation-induced microglia activation, we knocked down MyD88 by siRNAs (Supplementary Fig. 12d). We observed that downregulation of MyD88 did not affect LPS-induced activation of caspase-3 and caspase-8 (Supplementary Fig. 12e, f), suggesting that TLR4 ligation-induced microglia activation did not act through recruitment of the MyD88/FADD/caspase-8, as is the case for TLR2 ligation-induced apoptosis[26,27].

Microglia activation can be regulated by autocrine signalling of TNF-α by the TNF receptor 1 (TNFR1) secreted by LPS-stimulated microglia[28]. To examine whether we were observing a secondary effect, induced by autocrine signalling of TNF-α excreted by LPS-stimulated microglia cells, we investigated the effect of neutralizing TNF-α receptor antibodies on LPS-induced caspase-3 and caspase-8 activities[28] (Supplementary Fig. 13a-d). Selectively blocking the TNF-α receptor resulted in decreased LPS-induced iNOS expression at 24 h, thus confirming the TNF-α positive feedback loop exerted on microglia cells (Supplementary Fig. 13e). However, it failed to affect the LPS-induced caspase-3 and caspase-8 activities (Supplementary Fig. 13a–d). Thus, this experiment demonstrates that the caspase-signalling pathway is directly activated, independently of TNF-α receptor stimulation, upon LPS treatment.

Cleaved PKC-δ mediates effect of caspases

We then explored how activated caspase-3/7 interacts with the IKK/NF-κB pathway during LPS-induced microglia activation. PKC-δ has been reported to regulate NF-κB activation through the IKK complexes and phosphorylation of the NF-κB inhibitor IκB[29,30]. Interestingly, PKC-δ can be cleaved by caspases to generate a 40 kDa catalytically active fragment[31]. Because the levels of this protein are quite low in this cell line, we decided to overexpress PKC-δ to detect the cleaved form (Supplementary Fig. 14a). We found that LPS treatment of BV2 microglia cells promoted cleavage of PKC-δ into its 40-kDa active fragment (Fig. 4a, b). We found that the caspase inhibitor DEVD-fmk or selective siRNA knockdown of capase-3 or caspase-7 reduced LPS-induced PKC-δ activation. Therefore, in this context, PKC-δ activation is dependent on DEVD-ase activity (Fig. 4c, d). We then obtained further

evidence supporting that PKC-δ activation is important in microglia activation. First, we found that the PKC-δ inhibitor rottlerin inhibits LPS-induced iNOS expression in microglia (Fig. 4e, f). Second, if we overexpressed in microglia cells a caspase-uncleavable mutant of PKC-δ, we observed a decrease of LPS-induced iNOS expression (Fig. 4g, h). By contrast, when we overexpressed PKC-δ, we further enhanced the activating effect of LPS on microglia (Supplementary Fig. 14b). Also, we observed a marked decrease of all cytokines in primary microglial cell cultures after being challenged with rottlerin (Fig. 4i). Taken together, our experiments demonstrate that the caspase-3/7-dependent activation of microglia by LPS is mediated through PKC-δ.

Role of the caspase signalling *in vivo*

To examine the physiological relevance of these findings, we performed *in vivo* experiments and injected LPS into the rat substantia nigra. The substantia nigra is known to exhibit a strong inflammatory response upon LPS challenge[4]. At 24 h after injection (Fig. 5a–c), we observed a strong induction of caspase-8 and activation of caspase-3, which was mostly confined to reactive microglia in the mesencephalon on the injected side (Fig. 5b, c). By contrast, microglia were quiescent in the contralateral control midbrain (Fig. 5a). To study if caspase-3/7 activation is important for microglia activation also *in vivo*, we co-injected DEVD-fmk with LPS. Twenty-four hours later, we found that caspase-3/7 inhibition prevented LPS-induced microglia activation (Supplementary Fig. 15a, b). At the molecular level, we observed that the DEVD-fmk treatment mitigated LPS-induced expression of cytokines and pro-inflammatory molecules including iNOS, TNF-α and IL-1β (Fig. 5d). As a final step, we examined if our earlier *in vitro* findings about the role of PKC-δ in caspase-mediated microglia activation are relevant *in vivo*. We injected LPS intranigrally with IETD-fmk, DEVD-fmk or rottlerin. Twenty-four hours later, we dissected the substantia nigra and measured the iNOS protein level. In agreement with our *in vitro* data, we found that inhibition of caspase-8, caspase-3/7 or PKC-δ activities significantly decreased LPS-induced iNOS expression (Fig. 5e, f).

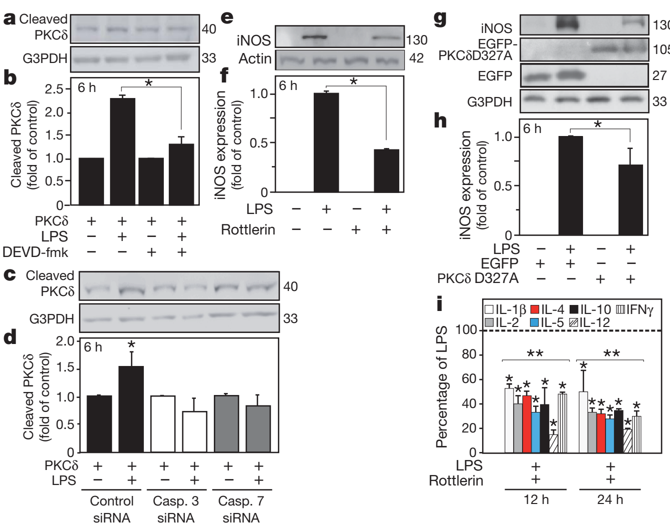


Figure 4 | Caspase-3/7 regulates microglia activation through the PKC-δ pathway. Caspase-3/7 inhibition by DEVD-fmk (**a**, **b**) or by siRNA knockdown (**c**, **d**) prevents LPS-induced PKC-δ cleavage/activation in BV2 microglia cells. PKC-δ inhibition by rottlerin mostly abolishes LPS-induced activation of iNOS (**e**, **f**). Overexpression of PKC-δ D327A, a caspase uncleavable PKC-δ mutant, has an inhibitory effect on LPS-induced iNOS expression (**g**, **h**). Rottlerin treatment reduces LPS-induced cytokine expression at 12 and 24 h in primary microglia cell culture (**i**). Data are expressed as mean ± s.e.m. ($n = 3$) and ± s.d. ($n = 4$) in **i**. *$P < 0.05$. In **i**, ** denotes statistically significant difference between treatments $P < 0.001$, concerning time and treatment.

To substantiate further our observations in the rat LPS model, we monitored the activation of microglia in the MPTP-lesion mouse model of PD. Consistent with earlier work, mice injected systemically with MPTP showed a strong microglial activation and a robust neurotoxicity in ventral mesencephalic dopamine neurons[32]. Intranigral vehicle injections in sham animals (saline containing 1% DMSO) greatly increased the density of reactive microglia, as evidenced by Iba1 immunohistochemistry (Supplementary Fig. 15c, d). Consequently, we quantified the numbers of reactive and resting microglia in Iba1-immunostained sections covering injection and non-injection sites for each experimental condition (Fig. 5g, h and Supplementary Fig. 15c, d). As expected, the density of reactive microglia at the non-injection site was highest in the MPTP group and dramatically lower in the sham (no MPTP) group. Interestingly, caspase 8 inhibition significantly prevented the MPTP-induced microglia activation and the MPTP-induced reduction in the density of resting microglia (Fig. 5g, h). Similar results of IETD-fmk were observed at the injection site (Supplementary Fig. 14c, d). As expected, MPTP injections severely reduced the integrity of nigro-striatal dopaminergic terminals (Supplementary Fig. 15e–g). Intranigral IETD-fmk injection induced a modest, but significant, protection against MPTP-induced toxicity as demonstrated by densitometric analysis of surviving striatal dopaminergic terminals and stereological cell counts in the substantia nigra (Supplementary Fig. 15e–g).

Caspase activation in human brains

PD and AD are known to be associated with neuroinflammation and the presence of activated microglia[2,33]. We investigated whether caspase-3 and caspase-8 are activated in microglia of individuals with PD and AD. We analysed expression of cleaved caspase-3, cleaved caspase-8 and the microglia marker CD68 in post-mortem brain samples from individuals with PD and AD (clinically as well as neuropathologically diagnosed) and age- and gender-matched healthy control brains. We detected significant cytoplasmic expression of both active caspase-3 and active caspase-8 in the PD ventral mesencephalon and in the AD frontal cortex, compared with controls (Fig. 6a–d and Supplementary Figs 16–19). The activated caspases and the microglial marker CD68 were largely co-localized, indicating that caspase-8 and caspase-3 are activated mainly in microglia in PD and AD (Fig. 6 a–d and Supplementary Fig. 20a, b).

Discussion

In this study, we uncover a completely novel role for caspase-8 and caspase-3/7 in the control of microglia and brain inflammation. We show that stimulation of microglia with the pro-inflammatory stimuli LPS, LTA, PamC3sk4 and IFN-γ triggers caspase-3/7 activation, without causing cell death. Caspases are proteases essential for apoptosis and inflammation. Caspase-1 is already known as the prototypical inflammatory caspase, required for the maturation of pro-inflammatory cytokines. By comparison, the initiator caspase-8 and effector pro-apoptotic caspase-3 and caspase-7 are considered crucial in the intracellular death machinery. Unexpectedly, we showed that caspase-3/7-dependent DEVD-ase activity controls LPS-induced microglia activation. We demonstrated that inhibition of the caspase-3/7 pathway effectively block microglia activation. For example, we observed fewer microglia with an activated phenotype in the presence of the caspase inhibitor DEVD-fmk and inhibition of the downstream IKK/NF-κB pathways. We also reveal that caspase-3/7 activates the IKK/NF-κB pathways through processing and activation of PKC-δ. Furthermore, we found that microglia exposed to the pro-inflammatory agent LPS failed to be toxic to neighbouring neurons when we inhibited caspase-3/7 chemically or by siRNA gene silencing. Importantly, we provide compelling evidence that active caspases 3 and 8 are expressed within reactive microglia in the ventral mesencephalon and frontal cortex of individuals with PD and AD, respectively.

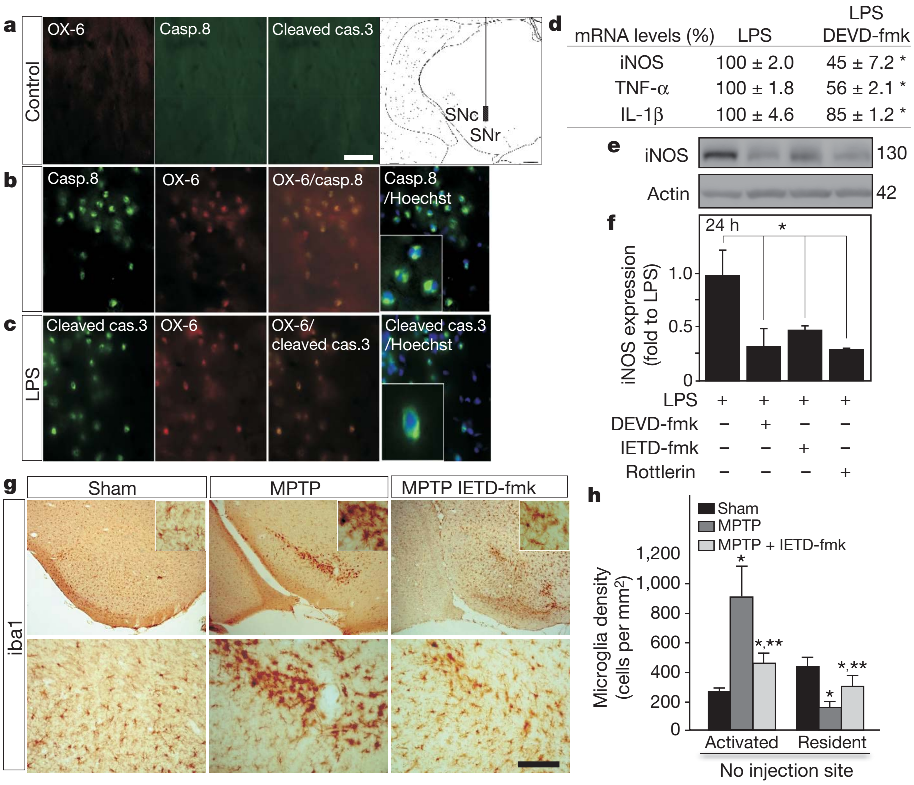


Figure 5 | *In vivo* inhibition of the caspase-dependent pathways prevents microglia activation. *In vivo* injection of LPS in rat substantia nigra (insert in **a**) induced activation of caspase-8 and caspase-3 in the ventral mesencephalon, which is mostly localized in OX-6-labelled reactive microglia (**b**, **c**) compared with control (**a**). There is a null coincidence between Hoechst-labelled nuclei and caspase8/cleaved caspase 3 immunoreactivity in OX-6-labelled microglia (**b**, **c**). *In vivo* DEVD-fmk significantly reduced intranigral LPS-induced microglia activation as evaluated by mRNA levels of proinflammatory molecules (**d**). *In vivo* inhibition of caspase-8, caspase-3/7 or PKC-δ, using IETD-fmk, DEVD-fmk and Rottlerin respectively, predominantly abolished the intranigral induction of iNOS upon LPS treatment (**e**, **f**). *In vivo* acute intraperitoneal injections of MPTP in mice highly increased the density of reactive microglia in substantia nigra (**g**, **h**). Intranigral IETD-fmk injections robustly prevented the MPTP-induced microglia activation and decrease of the resident microglia density at the non-injection site (**g**, **h**). Data are expressed as mean ± s.e.m. ($n = 3$). SNr, substantia nigra pars reticulata; SNc, Substantia nigra pars compacta. Scale bars: **a**–**c**, 25 μm; **d**, 500 μm for low magnification photographs; **j**, 400 and 100 μm for low- and high-magnification photographs. $^*P < 0.05$; ** denotes statistically significant difference between MPTP in **h**, and MPTP+IETD-fmk treatments.

To summarize, we present new non-apoptotic functions for caspases-8, -3 and -7 and show that they can have a pivotal role in inflammation of the central nervous system (Supplementary Fig. 1). Brain inflammation is a typical feature of neurodegenerative diseases[1,2,34] and is a prominent sequel of many acute forms of brain injury (for example, trauma, encephalitis and stroke)[35,36]. Under certain circumstances, neuroinflammation is known to promote neuronal death[37]. Accordingly, previous studies have shown that anti-inflammatory treatment can reduce neurodegeneration. Our discovery that the caspases-8, -3 and -7 cascade can promote neuroinflammation through IKK/NF-κB and PKC-δ, with the development of nanocarriers that allow the inhibitor to cross the blood–brain barrier, opens up new molecular targets for anti-inflammatory drugs[38]. Considering that these caspases also regulate apoptotic neuronal death, our results should revitalize interest in caspase inhibitors as potential therapeutic agents in disorders of the central nervous system[39–41].

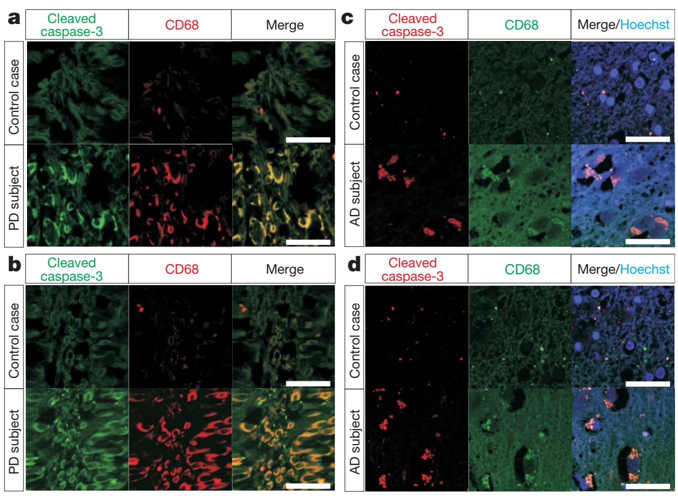


Figure 6 | Activation of caspase-3 and caspase-8 in microglia in brain from individuals with PD and AD. Double immunolabelling and confocal imaging analysis show significant activation of caspase-8 and caspase-3, mostly localized in CD68-labelled microglia, in the ventral mesencephalon from patients with PD (**a**, **b**) and in the frontal cortex of patients with AD (**c**, **d**) compared with age- and gender-matched healthy controls. Confocal images of immunostaining with cleaved caspase-3 (**a**, **c**) or cleaved caspase-8 (**b**, **d**), together with CD68 antibodies, are depicted. Merge from a control and an individual with AD illustrates the absence of cleaved caspase 3/caspase staining in Hoechst-labelled nuclei in reactive microglia (**c**, **d**). Scale bar, 10 μm.

METHODS SUMMARY

Microglial BV2 and dopaminergic MN9D murine cells were cultured as described[42,43]. Experiments were performed in reduced 5% FCS media. Transfection of BV2 cells used Lipofectamine 2000 (Invitrogen). Non-targeting control, caspase-3, caspase-7 and caspase-8 siRNA were obtained from Dharmacon. Single and co-cultures were exposed to 1 μg ml^{-1} LPS. Primary microglial cells were prepared from postnatal day 1–3 mouse brain using a previously described protocol[44]. Cytokine content was assayed using the Mouse TH1/TH2 9-PlexTissue Culture Kit (Meso-scale Discovery). DEVD-fmk (20 μM), IETD-fmk (20 μM), YVAD-fmk (20 μM), rottlerin (2 μM) or TNF R1 mouse Ab (20 ng ml^{-1}) were added to the media 1 h before LPS treatment. Under approved protocols, male albino Wistar rats (230–250 g) were intranigrally injected with 2 μg LPS alone or in combination with 0.75 nmol DEVD-fmk or 0.75 nmol IETD-fmk. Rottlerin (20 mg kg^{-1}) was administered intraperitoneally. Twenty-four hours after surgery, rats were killed and brains processed for analysis. Male C57BL mice were intranigrally injected with vehicle or 0.75 nmol IETD-fmk. Twelve hours later, animals were treated with four injections of MPTP (16 mg kg^{-1}) at 2-h intervals. Four days after the last injection, animals were killed and brains processed for analysis. Paraffin-embedded archival tissue blocks from autopsy on three individuals with PD and three individuals with AD were obtained from the Department of Pathology, Lund University Hospital, Sweden. Age-matched controls (cardiac arrest victims, no brain disease) were analysed at the same time. Histological, immunological and fluorescence-activated cell sorting (FACS) analyses using antibodies listed in Supplementary Table 1 were performed using standard procedures[42,45–47]. The subcellular localization of cleaved caspase-3, cleaved caspase-8 or NF-κB p65 protein was determined by confocal microscopy. Quantitative PCR with primers listed in Supplementary Table 2 were performed using SensiMixPlus SYBR (Quantace). MTT, caspase activity assay, apoptosis quantification and FACS analysis have been previously described[45]. Statistical evaluations were performed by one- or two-way analysis of variance with Bonferroni post hoc tests and Kolmogorov–Smirnov test.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 October; accepted 23 December 2010.

Published online 9 March 2011.

1. Hanisch, U. K. & Kettenmann, H. Microglia: active sensor and versatile effector cells in the normal and pathologic brain. *Nature Neurosci.* **10,** 1387–1394 (2007).
2. Block, M. L., Zecca, L. & Hong, J. S. Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. *Nature Rev. Neurosci.* **8,** 57–69 (2007).
3. Chao, C. C., Hu, S., Molitor, T. W., Shaskan, E. G. & Peterson, P. K. Activated microglia mediate neuronal cell injury via a nitric oxide mechanism. *J. Immunol.* **149,** 2736–2741 (1992).
4. Castano, A., Herrera, A. J., Cano, J. & Machado, A. Lipopolysaccharide intranigral injection induces inflammatory reaction and damage in nigrostriatal dopaminergic system. *J. Neurochem.* **70,** 1584–1592 (1998).
5. Saijo, K. *et al.* A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death. *Cell* **137,** 47–59 (2009).
6. Zhao, J. *et al.* IRF-8/interferon (IFN) consensus sequence-binding protein is involved in Toll-like receptor (TLR) signaling and contributes to the cross-talk between TLR and IFN-gamma signaling pathways. *J. Biol. Chem.* **281,** 10073–10080 (2006).
7. Car, B. D. *et al.* Interferon gamma receptor deficient mice are resistant to endotoxic shock. *J. Exp. Med.* **179,** 1437–1444 (1994).
8. Jin, J. J., Kim, H. D., Maxwell, J. A., Li, L. & Fukuchi, K. Toll-like receptor 4-dependent upregulation of cytokines in a transgenic mouse model of Alzheimer's disease. *J. Neuroinflammation* **5,** 23 (2008).
9. Balistreri, C. R. *et al.* Association between the polymorphisms of TLR4 and CD14 genes and Alzheimer's disease. *Curr. Pharm. Des.* **14,** 2672–2677 (2008).
10. Walter, S. *et al.* Role of the toll-like receptor 4 in neuroinflammation in Alzheimer's disease. *Cell. Physiol. Biochem.* **20,** 947–956 (2007).
11. Nicholson, D. W. *et al.* Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis. *Nature* **376,** 37–43 (1995).
12. Cohen, G. M. Caspases: the executioners of apoptosis. *Biochem. J.* **326,** 1–16 (1997).
13. Keller, M., Ruegg, A., Werner, S. & Beer, H. D. Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132,** 818–831 (2008).
14. Schulz, J. B. *et al.* Extended therapeutic window for caspase inhibition and synergy with MK-801 in the treatment of cerebral histotoxic hypoxia. *Cell Death Differ.* **5,** 847–857 (1998).
15. Braun, J. S. *et al.* Neuroprotection by a caspase inhibitor in acute bacterial meningitis. *Nature Med.* **5,** 298–302 (1999).
16. Cutillas, B., Espejo, M., Gil, J., Ferrer, I. & Ambrosio, S. Caspase inhibition protects nigral neurons against 6-OHDA-induced retrograde degeneration. *Neuroreport* **10,** 2605–2608 (1999).
17. Depino, A. M. *et al.* Microglial activation with atypical proinflammatory cytokine expression in a rat model of Parkinson's disease. *Eur. J. Neurosci.* **18,** 2731–2742 (2003).
18. Kawai, T. & Akira, S. Signaling to NF-κB by Toll-like receptors. *Trends Mol. Med.* **13,** 460–469 (2007).
19. Gibbons, H. M. & Dragunow, M. Microglia induce neural cell death via a proximity-dependent mechanism involving nitric oxide. *Brain Res.* **1084,** 1–15 (2006).
20. Schumann, R. R. *et al.* Lipopolysaccharide activates caspase-1 (interleukin-1-converting enzyme) in cultured monocytic and endothelial cells. *Blood* **91,** 577–584 (1998).
21. Li, P. *et al.* Mice deficient in IL-1β-converting enzyme are defective in production of mature IL-1β and resistant to endotoxic shock. *Cell* **80,** 401–411 (1995).
22. Friedlander, R. M. *et al.* Expression of a dominant negative mutant of interleukin-1β converting enzyme in transgenic mice prevents neuronal cell death induced by trophic factor withdrawal and ischemic brain injury. *J. Exp. Med.* **185,** 933–940 (1997).
23. Fernandes-Alnemri, T. *et al. In vitro* activation of CPP32 and Mch3 by Mch4, a novel human apoptotic cysteine protease containing two FADD-like domains. *Proc. Natl Acad. Sci. USA* **93,** 7464–7469 (1996).
24. Nunez, G., Benedict, M. A., Hu, Y. & Inohara, N. Caspases: the proteases of the apoptotic pathway. *Oncogene* **17,** 3237–3245 (1998).
25. Slee, E. A., Adrain, C. & Martin, S. J. Serial killers: ordering caspase activation events in apoptosis. *Cell Death Differ.* **6,** 1067–1074 (1999).
26. Aliprantis, A. O., Yang, R. B., Weiss, D. S., Godowski, P. & Zychlinsky, A. The apoptotic signaling pathway activated by Toll-like receptor-2. *EMBO J.* **19,** 3325–3336 (2000).
27. Jung, D. Y. *et al.* TLR4, but not TLR2, signals autoregulatory apoptosis of cultured microglia: a critical role of IFN-beta as a decision maker. *J. Immunol.* **174,** 6467–6476 (2005).
28. Kuno, R. *et al.* Autocrine activation of microglia by tumor necrosis factor-alpha. *J. Neuroimmunol.* **162,** 89–96 (2005).
29. Storz, P., Doppler, H. & Toker, A. Protein kinase Cdelta selectively regulates protein kinase D-dependent activation of NF-κB in oxidative stress signaling. *Mol. Cell. Biol.* **24,** 2614–2626 (2004).
30. Vancurova, I., Miskolci, V. & Davidson, D. NF-κB activation in tumor necrosis factor α-stimulated neutrophils is mediated by protein kinase Cδ. Correlation to nuclear IκBα. *J. Biol. Chem.* **276,** 19746–19752 (2001).
31. Reyland, M. E., Anderson, S. M., Matassa, A. A., Barzen, K. A. & Quissell, D. O. Protein kinase Cδ is essential for etoposide-induced apoptosis in salivary gland acinar cells. *J. Biol. Chem.* **274,** 19115–19123 (1999).
32. Czlonkowska, A., Kohutnicka, M., Kurkowska-Jastrzebska, I. & Czlonkowski, A. Microglial reaction in MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) induced Parkinson's disease mice model. *Neurodegeneration* **5,** 137–143 (1996).
33. Aarli, J. A. Role of cytokines in neurological disorders. *Curr. Med. Chem.* **10,** 1931–1937 (2003).
34. Gonzalez-Scarano, F. & Baltuch, G. Microglia as mediators of inflammatory and degenerative diseases. *Annu. Rev. Neurosci.* **22,** 219–240 (1999).
35. Jordan, J., Segura, T., Brea, D., Galindo, M. F. & Castillo, J. Inflammation as therapeutic objective in stroke. *Curr. Pharm. Des.* **14,** 3549–3564 (2008).
36. Lenzlinger, P. M., Morganti-Kossmann, M. C., Laurer, H. L. & McIntosh, T. K. The duality of the inflammatory response to traumatic brain injury. *Mol. Neurobiol.* **24,** 169–181 (2001).
37. Allan, S. M. & Rothwell, N. J. Inflammation in central nervous system injury. *Phil. Trans. R. Soc. Lond. B* **358,** 1669–1677 (2003).
38. Karatas, H. *et al.* A nanomedicine transports a peptide caspase-3 inhibitor across the blood-brain barrier and provides neuroprotection. *J. Neurosci.* **29,** 13761–13769 (2009).
39. Bilsland, J. & Harper, S. Caspases and neuroprotection. *Curr. Opin. Investig. Drugs* **3,** 1745–1752 (2002).
40. Friedlander, R. M. Apoptosis and caspases in neurodegenerative diseases. *N. Engl. J. Med.* **348,** 1365–1375 (2003).
41. Le, D. A. *et al.* Caspase activation and neuroprotection in caspase-3-deficient mice after *in vivo* cerebral ischemia and *in vitro* oxygen glucose deprivation. *Proc. Natl Acad. Sci. USA* **99,** 15188–15193 (2002).
42. Joseph, B. *et al.* p57(Kip2) cooperates with Nurr1 in developing dopamine cells. *Proc. Natl Acad. Sci. USA* **100,** 15619–15624 (2003).
43. Bocchini, V. *et al.* An immortalized cell line expresses properties of activated microglial cells. *J. Neurosci. Res.* **31,** 616–621 (1992).
44. Giulian, D. & Baker, T. J. Characterization of ameboid microglia isolated from developing mammalian brain. *J. Neurosci.* **6,** 2163–2178 (1986).
45. Li, J. Y. *et al.* Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature Med.* **14,** 501–503 (2008).
46. Joseph, B. *et al.* Mitochondrial dysfunction is an essential step for killing of non-small cell lung carcinomas resistant to conventional treatment. *Oncogene* **21,** 65–77 (2002).
47. Rite, I., Machado, A., Cano, J. & Venero, J. L. Blood-brain barrier disruption induces *in vivo* degeneration of nigral dopaminergic neurons. *J. Neurochem.* **101,** 1567–1582 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Gorman, O. Hermanson, M. Malewicz, S. Orrenius, T. Panaretakis and B. Zhivotovsky for discussion, and L. Hjortsberg, M. Reyland and S. Ceccatelli for providing us with reagents. M. Carballo, JL. Ribas, A. Fernández and B. Haraldsson provided qualified technical support. This work has been supported by grants from the Spanish Ministerio de Ciencia y Tecnología (SAF2006-04119 and 2009-13778), the Swedish Research Council, the Parkinson Foundation of Sweden, the Swedish Alzheimer Foundation and the Swedish Cancer Society. M.A.B., T.D. and P.B. are members of Neurofortis and Bagadilico, both of which are research environments sponsored by the Swedish Research Council.

Author Contributions M.A.B. performed all the experiments except as otherwise noted. qPCR was performed by A.G.-Q. and E.K. J.L.V. and J.C. collaborated in doing surgery and further dissecting the animal brains. M.A.B. and T.D. performed primary cell culture experiments and cytokine analysis. E.K. collaborated in performing the caspase activity assay. B.J. and E.K. collaborated in performing FACS. B.J. collaborated also in the confocal imaging analysis. E.E. did the neuropathology of the individuals with PD and AD and the controls. A.P. prepared tissue and participated in the morphological assessment of human brain specimens. N.H. and P.B. were involved in study design. M.A.B., J.L.V. and B.J. designed the study, analysed and interpreted the data. All authors discussed the results and commented on or edited the manuscript. The first draft of the paper was written by B.J. J.L.V. and B.J. share senior authorship of the paper. T.D and E.K. share second authorship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.J. (bertrand.joseph@ki.se) or J.L.V. (jlvenero@us.es).

METHODS

Reagents. LPS (from *Escherichia coli*, serotype 026:B6; Sigma), staurosporine (STS; Sigma), MPTP (Sigma), agonist anti-Fas monoclonal antibody clone CH11 (MBL), agonist anti-TNF-α (R&D Systems), the caspase-3/7 inhibitor benzyloxycarbonyl-Asp(OMe)-Glu(OMe)-Val-Asp(OMe)-fluoromethylketone (DEVD-fmk), the caspase-8 inhibitor benzyloxycarbonyl-Ile-Glu(OMe)-Thr-Asp(OMe)-fluoromethylketone (IETD-fmk), the caspase-6 inhibitor benzyloxycarbonyl-Val-Glu(OMe)-Ile-Asp(OME)-fluoromethylketone (VEID-fmk), the caspase-1 inhibitor benzyloxycarbonyl-Tyr-Val-Ala-Asp(OMe)-fluoromethylketone (YVAD-fmk; MP Biomedicals) and the PKC-δ inhibitor Rottlerin (Calbiochem) were used in this study. Plasmids encoding PKC-δ and EGFP-PKC-δ D327A were gifts of L. Hjörstberg and M. Reyland respectively. ON-TARGET plus SMARTpools siRNAs were purchased from Dharmacon (Supplementary Table 3).

Animals and surgery. Animals used in this study were obtained from the Center of Production and Animal Experimentation (Espartinas, Spain) and NMRI (Charles River, Germany). Experiments were performed in accordance with the Guidelines of the European Union Council (86/609/EU), following Spanish and Swedish regulations for the use of laboratory animals and approved by the Scientific Committee of the University of Seville and Lund University. Intranigral injections were made 5.8 mm anterior, 2.0 mm lateral and 8.0 mm ventral to the bregma in rat, and 3.1 mm anterior, ± 1.2 mm lateral and 5.1 mm ventral from bregma in mouse.

Co-culture and neuronal cell death assay. Microglial BV2 and dopaminergic MN9D murine cells were cultured as described[42,43]. MN9D dopaminergic neuronal cells were stained with CellTracker Green CMFDA (Invitrogen) before BV2 microglia cells transfected with caspase-3 siRNA, caspase-7 siRNA or non-targeting siRNA were plated on them. After 24 h, cells were treated with LPS and incubated for an additional 24 h. They were then fixed with 4% paraformaldehyde and stained with 0.1 mg ml^{-1} Hoechst for quantification of neuronal cell death.

Primary cultures. Primary mouse microglial cells were plated for at least 48 h before the experiments (ethical permit M302-09). Cells were pretreated with inhibitors (20 μM DEVD-fmk, 20 μM IETD-fmk and 2 μM Rottlerin 1 h before with 100 ng ml^{-1} LPS treatment) [44].

Human brain. Human brain tissues from patients with PD of 5, 9 and 15 years' duration, respectively, and age-matched controls were used in this study[45]. The region investigated was the anterior mesencephalon covering the substantia nigra. Frontal cortex from patients with AD of 4, 10 and 14 years' duration and controls were also used (Regional Ethical Review Board, Lund (Sweden) 2009-646/2010-25). The patients with dissimilar disease duration exhibited different degrees of severity of brain disease, reflecting different stages of the degenerative process. All sections were stained with haematoxylin–erythrosin and with antibodies against caspase 3, 8 and microglia, CD68 (see below). They were microscopically reviewed for verification of pathology. Before the investigation, the entire collection of brain sections, 15–20 per individual including the mesencephalic section, were subjected to a neuropathological whole-brain analysis for clinical diagnostic purposes, according to routine procedures at the Department of Pathology, Division of Neuropathology, Lund University Hospital. The project procedures involving human brain tissue were approved by the Regional Ethical Review Board in Lund, Sweden.

Immunofluorescence and laser scanning confocal microscopy. Paraformaldehyde-fixed cells were blocked in PBS/3% goat serum/0.3% triton X-100 and incubated with the indicated primary (4 °C, overnight) and AlexaFluor 488 or 594 conjugated anti-IgG used as secondary antibodies (room temperature, ~22 °C, 1 h; Molecular Probes). Nuclei were counterstained with DRAQ5^{TM} (1 μM, Alexis), Hoecsht (1 μg ml^{-1}, Molecular Probes) or DAPI (1 μg ml^{-1}, Molecular Probes). Alexa-555-conjugated cholera toxin B (CTB; 10 μg ml^{-1}; Molecular Probes) was used to stain lipid rafts on the plasma membrane. Protein subcellular localization was analysed under a Zeiss 510 META or Leica TCS-SP2 confocal laser scanning microscope. The nuclear translocation of the NF-κB p65 subunit was quantified using multicolour three-dimensional plug-in from Leica Confocal software and measured as the percentage of co-localization of DRAQ5^{TM} with NF-κB p65 inside the nucleus.

FACS analysis. Quantification of cells with cleaved caspase-3 and cleaved caspase-8 was performed with a FACSCalibur flow cytometer (Becton Dickinson) using standard procedures. Analysis of data was performed using Cell Quest software[46].

Immunohistology and immunohistochemistry. Rats and mice were perfused through the heart under deep anaesthesia with 4% paraformaldehyde/PBS, pH 7.4. Brains were removed, cryoprotected in sucrose and frozen in isopentane at −15 °C; serial coronal sections (25- and 30-μm sections, for rat and mice, respectively) covering the striatum and the substantia nigra were cut with a cryostat and mounted on gelatin-coated slides[47]. Sections were incubated with the indicated primary antibodies. After three washes, sections were incubated with biotinylated horse anti-mouse or goat anti-rabbit IgG (Vector) followed by an incubation with ExtrAvidin-Peroxidase solution (Sigma) and for immunofluorescence by a fluorescein isothiocyanate (FITC)-conjugated anti-rabbit and Texas Red anti-mouse antibody (Vector). The peroxidase was visualized with a standard diaminobenzidine/hydrogen reaction for 5 min. For paraffin-embedded human tissue material, sections (5 μm) were mounted on capillary glass slides (DAKO). Sections were microwaved pre-treated in 10 mM citrate buffer pH 6.0 for 15 min at 800 W for antigen retrieval. An automated immunostainer (TechMateTM 500 Plus, DAKO) was used for the staining procedure using DAKO ChemMate Kit Peroxidase/3-3′diaminobenzidine. The indicated primary antibodies were used.

Quantification of microglial population in animal models. Reactive and resident microglial cells were counted in LPS-injected rats and MPTP-exposed mice detected by Iba1 immunohistochemistry based on morphological features. For each animal, eight sections covering the entire antero-posterior ventral mesencephalon were analysed. For each section, four photographs were taken at ×20 magnification (two for each substantia nigra) and microglial cells were counted with the analySIS software.

Quantification of the striato-nigral dopaminergic system. Tyrosine hydroxylase immunohistochemistry in the striatum of mice intoxicated with MPTP was quantified using a computer-assisted software (analySIS). For the quantification, five striatal sections from each condition, processed under identical experimental conditions, were scanned at high resolution. The striatal region was delineated and its optical density measured based upon a calibrated grey scale. Quantification of tyrosine-hydroxylase-positive cells in the substantia nigra was performed according to a modified stereological approach using the Olympus CAST-Grid system. The area of the substantia nigra region was estimated using the principle of Cavallieri. All data were collected while blind to experimental treatment and expressed as number of neurons per substantia nigra.

doi:10.1038/nature09853

Crystal structure of inhibitor of κB kinase β

Guozhou Xu[1]*, Yu-Chih Lo[1]*, Qiubai Li[1], Gennaro Napolitano[2], Xuefeng Wu[2], Xuliang Jiang[3], Michel Dreano[4], Michael Karin[2] & Hao Wu[1]

Inhibitor of κB (IκB) kinase (IKK) phosphorylates IκB proteins, leading to their degradation and the liberation of nuclear factor κB for gene transcription. Here we report the crystal structure of IKKβ in complex with an inhibitor, at a resolution of 3.6 Å. The structure reveals a trimodular architecture comprising the kinase domain, a ubiquitin-like domain (ULD) and an elongated, α-helical scaffold/dimerization domain (SDD). Unexpectedly, the predicted leucine zipper and helix-loop-helix motifs do not form these structures but are part of the SDD. The ULD and SDD mediate a critical interaction with IκBα that restricts substrate specificity, and the ULD is also required for catalytic activity. The SDD mediates IKKβ dimerization, but dimerization per se is not important for maintaining IKKβ activity and instead is required for IKKβ activation. Other IKK family members, IKKα, TBK1 and IKK-i, may have a similar trimodular architecture and function.

Nuclear factor κB (NF-κB) transcription factors are master regulators of inflammatory, immune and apoptotic responses[1,2]. In the canonical pathway, NF-κB dimers are held in the cytoplasm through binding to IκB proteins, which mask their nuclear localization signals. When cells are stimulated by NF-κB inducers, IκB proteins are phosphorylated by the Ser/Thr-specific IKK, a modification that triggers their Lys-48-linked polyubiquitination and subsequent proteasomal degradation[3]. Freed NF-κB dimers enter the nucleus to regulate gene transcription[2]. In the non-canonical pathway, activated IKK phosphorylates the IκB-like domain in the NF-κB family member NF-κB2 (p49/p100)[3] (NFKB2 in humans). NF-κB signalling pathways are associated with a vast number of human diseases including inflammatory disorders and cancer, which renders IKK a potentially important therapeutic target[4] (see, for example, http://www.nf-kb.org).

IKK was originally purified from HeLa cells as a multiprotein complex that contains the kinase subunits IKKα (CHUK) and/or IKKβ (IKBKB), and the regulatory protein NEMO[5–11] (IKKγ, or IKBKG). IKKα and IKKβ both contain an amino-terminal kinase domain (KD), predicted leucine zipper (LZ) and helix–loop–helix (HLH) domains, and a carboxy-terminal NEMO-binding domain (Fig. 1a). IKKβ seems to have an additional ULD following the KD, which is not predicted in the corresponding region of IKKα[3]. The IKK-related kinases TBK1 (NAK) and IKK-i (IKK-ε, or IKBKE) seem to have a similar domain organization[12]. Whereas IKKβ mediates activation of the canonical NF-κB pathway in response to pro-inflammatory stimuli, IKKα has an indispensible role in non-canonical NF-κB signalling by phosphorylating NF-κB2 (p49/p100). Protein kinase assays suggest that IKKβ accounts for nearly all of the catalytic activity of the IKK holoenzyme towards IκBα[3,13] (NFKBIA).

The activation loop in both the IKKα and the IKKβ KD contains the mitogen-activated protein (MAP)-kinase kinase consensus motif SXXXS, where X is any amino acid[6–8,10] (Ser 177 and Ser 181 in human IKKβ). Some MAP-kinase kinase kinase family members, such as TGF-β-activated kinase 1 (MAP3K7, previously TAK1) and NF-κB-inducing kinase (MAP3K14, or NIK), were shown to phosphorylate IKKs[3,14,15]. IKKα and IKKβ can also undergo autophosphorylation and activation as a result of overexpression or signal-dependent NEMO clustering[10,16]. Ala substitutions of the activation-loop Ser residues prevent IKK activation whereas the phosphomimetic, double Glu mutation S177E/S181E (EE) of IKKβ renders it constitutively active[7,13].

Trimodular arrangement of IKKβ

We determined the crystal structure of *Xenopus laevis* IKKβ (ikbkb) EE (residues 4–675; Fig. 1a) in complex with either inhibitor Cmpd1 or inhibitor Cmpd2 (Supplementary Fig. 1) at resolutions of 4.0 and 3.6 Å in the $I4_122$ and $P1$ space groups, respectively (Supplementary Tables 1 and 2 and Supplementary Fig. 2). Eight IKKβ molecules in $P1$ and the single molecule in $I4_122$ are highly similar to each other (Supplementary Fig. 3 and Supplementary Table 3) and show conserved dimerization (see below). Our structural description will use the first dimer (chains A and B) in $P1$. The sequences of human and *Xenopus* IKKβ (henceforth hIKKβ and xIKKβ, respectively) have 74% identity with no gaps within the region of the structure; residue numbers designated for xIKKβ are also true for hIKKβ.

The IKKβ dimer structure resembles a pair of shears and has the overall dimensions of approximately 100 Å × 130 Å × 60 Å (Fig. 1b, c). It contains the KD (residues 16–307), the ULD (310–394) and the SDD (410–666) (Fig. 1a and Supplementary Fig. 4). The KD and the ULD form the 'handle' of the shears, and the SDD is the 'blade'. Both the KD and the ULD intimately associate with the SDD, suggesting that the KD does not function independently. Instead, the structure indicates that IKKβ is an integral trimodular unit.

The IKKβ KD in complex with either Cmpd1 or Cmpd2 has the typical bilobal kinase fold[17]. Although it has only 21.1% sequence identity with human ubiquitin, the ULD of IKKβ indeed has the ubiquitin fold (Fig. 1d). A major difference is the presence of an eight-residue insertion (373–380) that forms part of the loop connecting β-strands β4′ and β5′ in the ULD, which is mostly disordered. The hydrophobic surface patch of ubiquitin[18], which is often recognized by ubiquitin-binding proteins, is conserved in the ULD, with residues Val 318, Leu 354 and Leu 389 equivalent to Leu 8, Ile 44 and Val 70 of ubiquitin (Supplementary Fig. 5).

The SDD comprises six α-helices (α1s–α6s), of which α2s and α6s contain 70 and 77 residues, respectively. They twist together with a stretch of three shorter helices between them to fold as an elongated structural ensemble. The C lobe of the KD sits on the N-terminal end

[1]Department of Biochemistry, Weill Cornell Medical College, New York, New York 10021, USA. [2]Department of Pharmacology, University of California at San Diego, La Jolla, California 92093, USA. [3]EMD Serono Research Institute, Billerica, Massachusetts 01821, USA. [4]Merck Serono, Geneva 1211, Switzerland.
*These authors contributed equally to this work.

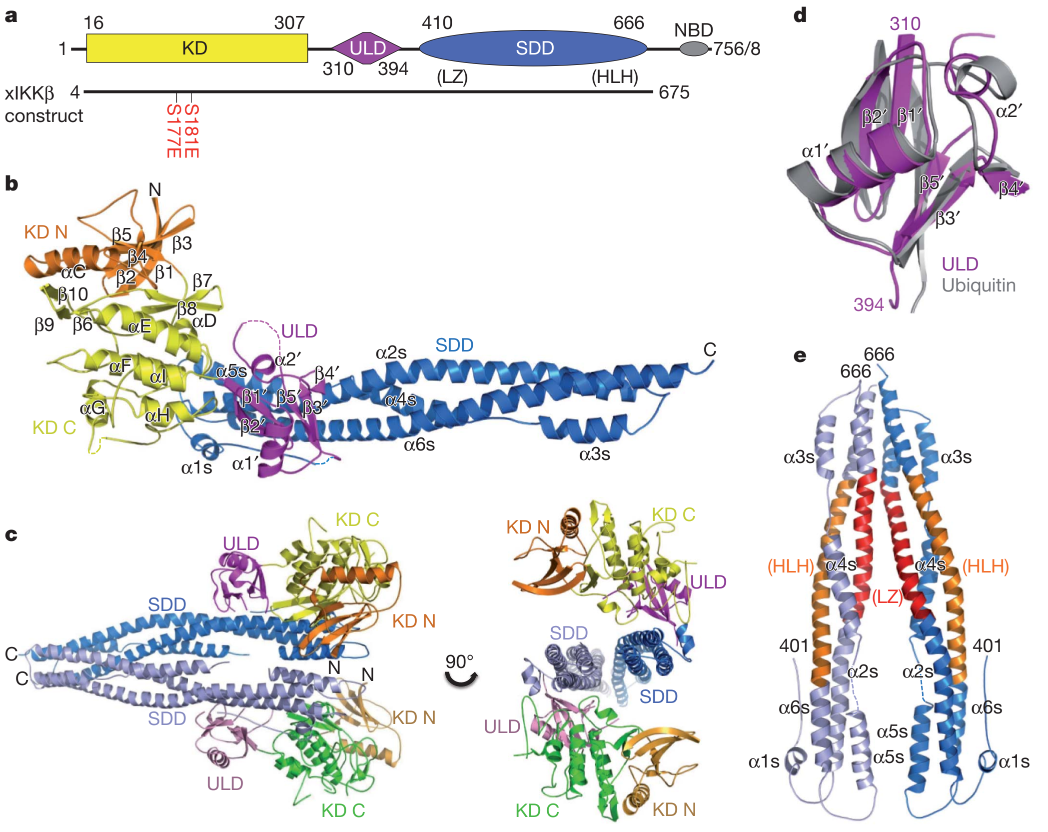


Figure 1 | Structure of xIKKβ. **a**, Linear representation of IKKβ showing the boundaries for the KD, the ULD and the SDD. Sequences of hIKKβ and xIKKβ are of 756 and 758 residues, respectively, and differ only at the most C-terminal region. The crystallized xIKKβ construct is shown. The previously designated LZ and HLH regions are shown in parentheses. **b**, Ribbon diagram of an xIKKβ protomer in the *P*1 crystal form. The N and C termini, KD N lobe (orange) and C lobe (yellow), ULD (magenta) and SDD (blue) are labelled. Secondary structural elements are labelled, with those in the ULD indicated with a prime and those in the SDD indicated with an 's'. **c**, Ribbon diagram of an xIKKβ dimer. **d**, Superposition of the ULD (magenta) with ubiquitin (grey). **e**, Ribbon diagram of an SDD dimer, showing locations of the previously designated LZ (red) and HLH (orange) regions.

of the SDD, and the ULD binds close to the middle of the SDD. Unexpectedly, formerly designated LZ (458–485) and HLH domains[7,8,10] (605–644) do not exist as such in the structure and are both part of the SDD (Fig. 1e). Most of the hydrophobic residues in the predicted LZ point inwards and are not available for mediating dimerization as previously proposed.

Structure of inhibitor-bound IKKβ KD

The inhibitor binds to the IKKβ KD at the hinge loop connecting the N and C lobes, a region that recognizes the adenine in ATP[19,20] and is often used for inhibitor binding[21–23] (Fig. 2a and Supplementary Fig. 6). The KD conformation is incompatible with that of an active kinase[17,24,25] (Fig. 2b, c). The activation segment, which begins from Asp 166-Leu 167-Gly 168 as the conserved DFG triad and extends to Ala 190-Pro 191-Glu 192 in the conserved APE motif, is fully ordered, including phosphomimetic residues Glu 177 and Glu 181 (Fig. 2b). However, the C-terminal anchor of the activation segment, including the APE motif, is completely out of place, with the result that essential interactions with the catalytic centre are lost (Fig. 2c). The gross conformation of the N-terminal anchor of the activation segment is preserved, with β9 paired with the β6 strand that precedes the catalytic loop. Part of the activation loop (175–177) contains an additional β-strand (β10), which sits between the lobes to form a three-stranded β-sheet with β9 and β6. The αC helix is tilted up and outwards (Fig. 2c) to make room for β10, which also disrupts its productive interactions with the DFG motif in active kinase structures.

Interactions among the KD, ULD and SDD

The KD, ULD and SDD interact mutually, with the ULD–SDD interaction being the most extensive, followed by the KD–SDD and KD–ULD

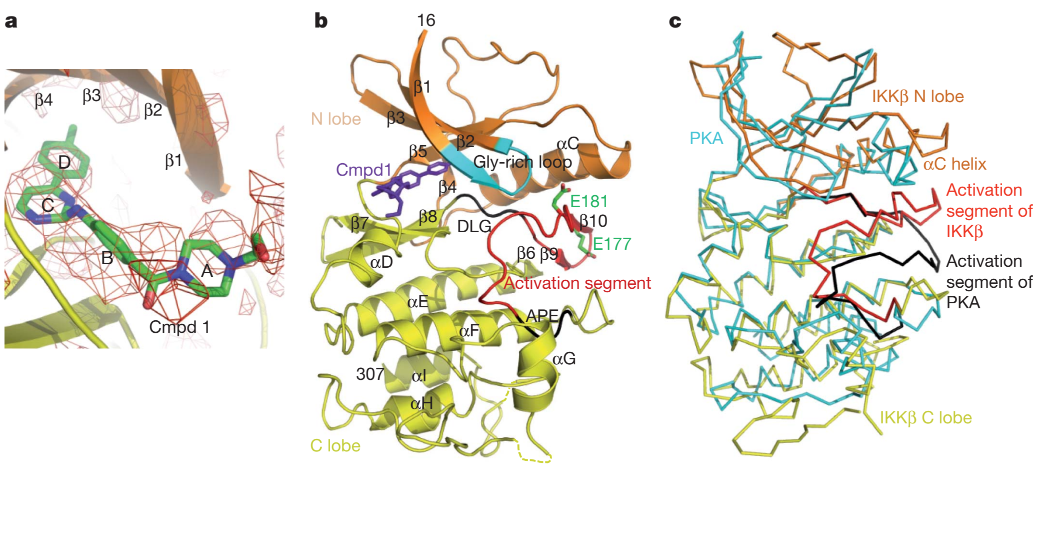


Figure 2 | Inhibitor-bound xIKKβ kinase domain. **a**, $F_o - F_c$ electron density map for Cmpd1 in the $I4_122$ structure, contoured at 2.0σ. Carbon, nitrogen and oxygen atoms are shown in green, blue and red, respectively. The four ring structures in Cmpd1 are labelled A, B, C and D, respectively. **b**, Structure of the xIKKβ KD. Gly-rich loop, cyan; activation segment, red (except that the DLG and APE sequences are in black); Cmpd1, purple. Side chains of phosphomimetic residues Glu 177 and Glu 181 are shown. **c**, Superposition of xIKKβ (orange and yellow) and protein kinase A (PKA, cyan; Protein Data Bank ID, 1ATP). The activation segments of xIKKβ and PKA are shown in red and black, respectively.

interactions. The ULD binds close to the middle of the SDD, contacting helices α2s and α6s (Fig. 3a). The interaction buries ~700 $Å^2$ of surface area per partner. There are substantial hydrophobic contributions, supplemented by electrostatic interactions. Residues Met 315, Met 317, Leu 354, Ile 387, Leu 389 and Phe 390 on one side of the ULD pack against Leu 612, Tyr 609, Leu 447 and the main chain of α2s of the SDD to form the central hydrophobic core of the interface. This hydrophobic patch of the ULD is immediately adjacent to and overlaps the conserved hydrophobic patch in the ubiquitin fold. Electrostatic interactions are observed between Glu 352 and Lys 619 and between Lys 394 and Asp 610. Additional interfacial residues include Ser 319 and Ser 357 of the ULD and Thr 453, Gln 449 and Val 616 of the SDD.

Consistent with an important role of the ULD in interacting with the SDD, mutations in residues that are not at the interface, P347A, L361A and Q351A, had minimal effects on NF-κB activation[26]. In contrast, mutation in a residue within a ULD surface loop that contacts the SDD, G358A (Fig. 3a), affected IKKβ-induced NF-κB activity[26]. It was also shown that Leu 353 is required for IKKβ activity[26]; however, Leu 353 is buried in the ULD core and the L353A mutation may have disrupted IKKβ structural integrity. Double substitutions of hIKKα and hIKKβ, which are equivalent to I608R/Y609P of the SDD of xIKKβ, did not affect their interaction with wild-type IKKβ but negatively impacted kinase activity[10]; Ile 608 is buried in the SDD core and Tyr 609 directly interacts with the ULD (Fig. 3a).

Like the ULD, the KD also makes contact with the N terminal end of the SDD (Fig. 3b), burying a surface area of ~650 $Å^2$ of each interface. The binding forces are mainly van der Waals in nature. Limited hydrophobic interactions are observed between residues Ala 252 and Val 253 of the KD and Tyr 423 of the SDD, and between Phe 111 of the KD and the hydrophobic portions of Arg 572, Arg 575 and Glu 576 of α5s of the SDD. The KD is linked to the ULD through a two-amino-acid linker between αI of the KD and β1′ of the ULD (Fig. 3c), burying only ~350 $Å^2$ of surface area. Side-chain contacts between Leu 311 of the ULD and Ile 268 of the KD are observed, and together form the small hydrophobic patch at the KD–ULD junction consisting also of Leu 269 and Ile 306 of the KD and Leu 309 of the linker. Together with the ionic interaction between Asp 373 of the ULD and Arg 123 of αE of the KD, these interactions may confer rigidity to the KD–ULD junction despite the small buried surface area.

Structural comparison with other kinase structures revealed a similarity between the locations of the SDD and ULD and those of several known docking sites for substrates and regulatory proteins[27]. In the crystal structure of the Ser/Thr kinase SRPK1 in complex with a docking motif in its substrate, ASF/SF2[27] (SRSF1), the peptide motif interacts with the kinase at the location of the SDD (Fig. 3d). In the structure of the TAK1 KD fused with the C-terminal region of its binding protein, TAB1[28], TAB1 interacts with the C lobe of the kinase at a position analogous to both the SDD and the ULD, presumably to activate the kinase (Fig. 3d).

ULD–SDD binds the IκBα C-terminal region

Previous studies have suggested that the ULD in TBK1 and IKK-i is involved in substrate recognition because its deletion impaired activity of the respective kinases and a glutathione *S*-transferase (GST)–ULD fusion protein interacted with the specific substrate, IRF3 or IRF7[29]. Because ULD deletion in IKKβ also abolished its activity[26], we proposed that the ULD may recognize its specific substrate, IκBα. However, we were surprised to find that GST–IκBα (1–317 and 54–317) pulled down only a minute amount of the ULD (Fig. 4a, lanes 9 and 13) and GST alone did not pull down any (Fig. 4a, lane 4), suggesting that the interaction of IκBα with the ULD is specific, but very weak. In contrast, GST–IκBα robustly pulled down full-length IKKβ or IKKβ lacking C-terminal NEMO-binding domain (Fig. 4a, lanes 15 and 16).

IκBα has an N-terminal region (1–54) that contains cognate phosphoacceptor sites at Ser 32 and S 36, followed by a C-terminal region (55–317) that contains multiple ankyrin repeats and the PEST region[30,31]. Strikingly, the N-terminal region of IκBα did not pull down any IKKβ constructs (Fig. 4a, lanes 5–8), and the C-terminal region of IκBα interacted robustly with full-length IKKβ as well as its ULD–SDD region (Fig. 4a, lanes 9–12), and very weakly with the ULD alone (Fig. 4a, lane 9). Further mapping on IκBα showed that both

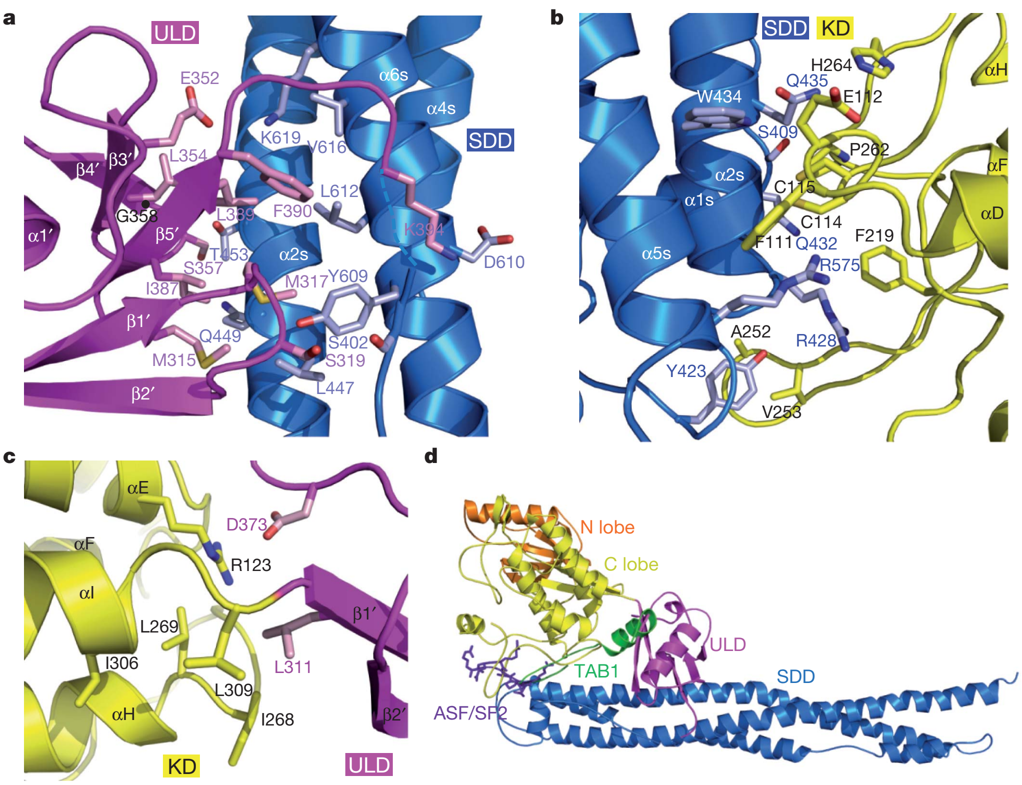


Figure 3 | Interactions among the KD, ULD and SDD. **a**, Interaction between the ULD (magenta) and SDD (blue). Important interfacial side chains are shown with nitrogen atoms in blue, oxygen atoms in red, sulphur atoms in orange and carbon atoms in either pink (ULD) or light blue (SDD). Gly 358 is marked with a black ball on the main chain. **b**, Interaction between the KD (yellow) and SDD (blue). **c**, Interaction between the KD (yellow) and ULD (magenta). **d**, Locations of the TAB1 peptide (green ribbon; PDB ID, 2EVA) and the ASF/SF2 peptide (purple stick model; PDB ID, 1WBP) relative to the IKKβ structure after superposition of the corresponding kinase domains.

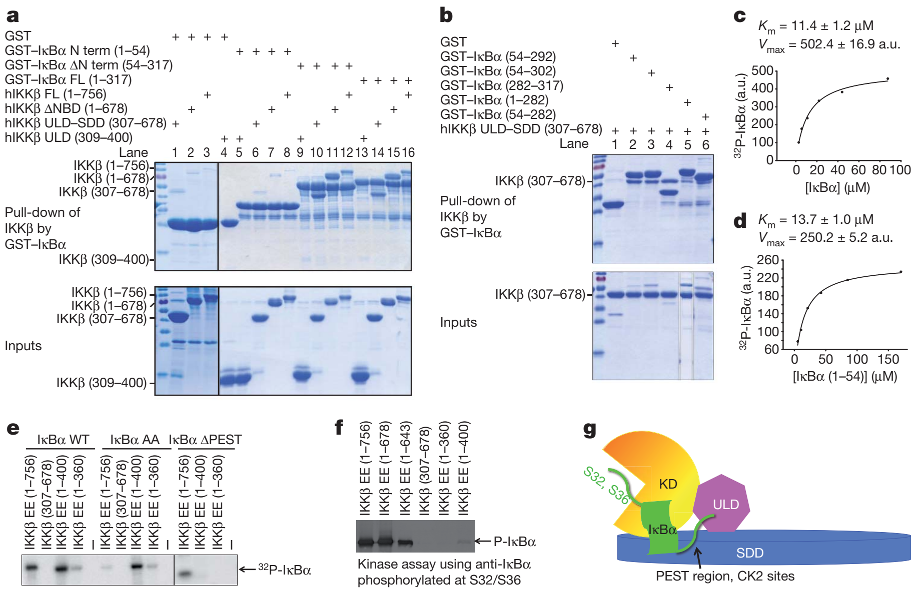


Figure 4 | ULD–SDD restricts IKKβ specificity and ULD is required for catalytic activity. **a**, **b**, Pull-down of hIKKβ constructs using GST–IκBα constructs, showing the reciprocal interaction between ULD–SDD of IKKβ and the C-terminal region of IκBα containing ankyrin repeats and the PEST region. FL, full length. **c**, **d**, Measurement of K_m and the relative V_{max} of IKKβ against full-length IκBα (**c**) and the N-terminal region of IκBα (1–54) (**d**). a.u., arbitrary units. **e**, Kinase assay of purified hIKKβ proteins against IκBα, its S32A/S36A mutant (AA) and its PEST-deletion construct (ΔPEST, 1–282) using [γ-^{32}P]ATP. WT, wild type. **f**, Kinase assay of purified hIKKβ proteins using antibody against IκBα phosphorylated at Ser 32 and Ser 36. **g**, Interaction between the SDD of IKKβ and the C-terminal region of IκBα may position the N-terminal cognate phosphorylation sites of IκBα at the active site of IKKβ. CK2, casein kinase 2.

ankyrin repeats (1–282 or 54–282) and the PEST region (282–317) interacted with IKKβ ULD–SDD (Fig. 4b, lanes 4–6). The extent of pull-down suggests that the PEST region contributes more to IKKβ interaction than do the ankyrin repeats. Despite trying multiple constructs, we could not obtain soluble protein for IKKβ SDD to test its interaction with IκBα. However, the pull-down data suggest a dominant role for SDD in IκBα interaction. In any case, it is clear that the mutual interaction between IKKβ and IκBα is mediated by their ULD–SDD and C-terminal domains, respectively.

ULD–SDD limits specificity and the ULD aids catalysis

The specific interaction between ULD–SDD of IKKβ and IκBα suggests that the Michaelis constant, K_m, of phosphorylation by IKKβ might be lower for full-length IκBα than for its N-terminal region (1–54) alone. We performed kinase kinetics analysis of IKKβ EE against the two different substrates. Unexpectedly, the measured apparent K_m values were 11.4 and 13.7 μM for full-length IκBα and the N-terminal region alone, respectively (Fig. 4c, d), suggesting that binding at the SDD, an exosite, does not alter the K_m of the reaction drastically. This could be due to opposing effects of the SDD–IκBα interaction, which increases substrate interaction but slows down product dissociation. The relative maximum velocity, V_{max}, values were 502.4 and 250.2, respectively, suggesting that the SDD–IκBα interaction moderately enhances catalysis.

Casein kinase 2 phosphorylates the C-terminal PEST region of IκBα around residues 283–299[30,31]. To determine whether the SDD–IκBα interaction restricts substrate specificity, we compared the kinase activity of purified IKKβ EE proteins against either IκBα or its S32A/S36A mutant (AA) using [γ-^{32}P]ATP (Fig. 4e). Although the KD–ULD (1–400) construct gave rise to a small amount of protein, it produced robust phosphorylation of wild-type IκBα, comparable to full-length IKKβ (1–756), suggesting that it is catalytically competent. Remarkably, KD–ULD effectively phosphorylated the AA mutant, which, in contrast, was a poor substrate for full-length IKKβ. The C-terminal PEST region seemed to harbour the major sites of phosphorylation by KD–ULD because a construct lacking PEST (1–282) was not significantly phosphorylated by KD–ULD but was phosphorylated by full-length IKKβ (Fig. 4e). In addition, when IκBα phosphorylation was detected by antibody against IκBα phosphorylated at Ser 32/Ser 36, only very weak phosphorylation was seen for the KD–ULD construct in comparison with full-length IKKβ (Fig. 4f). These experiments suggest that although the KD–ULD construct is active, it possesses an altered substrate specificity causing preferential phosphorylation of the C-terminal PEST sequence, consistent with a previous observation[32]. Hence, ULD–SDD seems to position IκBα specifically such that only its N-terminal region is accessible to the IKKβ catalytic pocket (Fig. 4g).

We expressed three kinase domain constructs, 1–301, 1–310 and 1–360, in insect cells and obtained only small amount of protein for the 1–360 construct. Kinase assay showed that IKKβ EE (1–360) had very low residual activity against IκBα or its AA mutant (Fig. 4e), suggesting that KD activity is compromised in the absence of the ULD. Without an isolated KD structure, we cannot deduce the molecular mechanism by which the ULD acts. However, in analogy to TAK1 activation by TAB1[28] (Fig. 3d), it may be speculated that this KD–ULD interaction allosterically potentiates kinase activity. Further kinase assay using antibody against IκBα phosphorylated at Ser 32/Ser 36 showed no detectable activity of the KD alone (Fig. 4f), confirming that the low residual activity may also be directed towards the C-terminal PEST region, similar to KD–ULD. Therefore, whereas ULD–SDD is critical for IKKβ specificity, ULD is required for its catalytic activity.

SDD mediates IKKβ dimerization

Full-length hIKKβ (1–756) and its KD–ULD–SDD region (1–678) are dimers in solution as determined by gel-filtration chromatography (Fig. 5a). In the *P*1 and $I4_122$ crystals, two conserved dimerization interfaces exist, one mediated by the SDD and the other mediated by the KD. Because hIKKβ (1–643) that truncates into the SDD is a monomer in solution (Fig. 5a), SDD-mediated dimerization (Fig. 5b) is probably of greater importance than KD-mediated dimerization.

SDD in an IKKβ dimer does not form extensive interactions along the entire length dimension of the helical bundle. Rather, interactions are mostly localized at the end of the bundle (Fig. 5b), distal from the KD and ULD and burying ~1,000 Å^2 of surface area of each monomer. The interaction is mostly hydrophobic. Residues that contribute significantly to dimerization include Gln 478, Lys 482, Phe 485, Ile 492, Lys 496, Ile 505, Gln 651, Leu 654, Trp 655, Leu 658 and Ile 660, with residues Leu 654, Trp 655 and Leu 658 from helix α6s burying the most surface area (Fig. 5b). This dimerization interface is entirely unexpected as it was predicted earlier that the LZ mediates IKKβ dimerization. The structure now explains that failed dimerization of the LZ-defective L462S/L469S mutant of IKKα[10] is due to the structural role of Leu 469, whose equivalent residue of IKKβ, Met 472, is buried in the SDD core. Superposition of four IKKβ dimers in *P*1

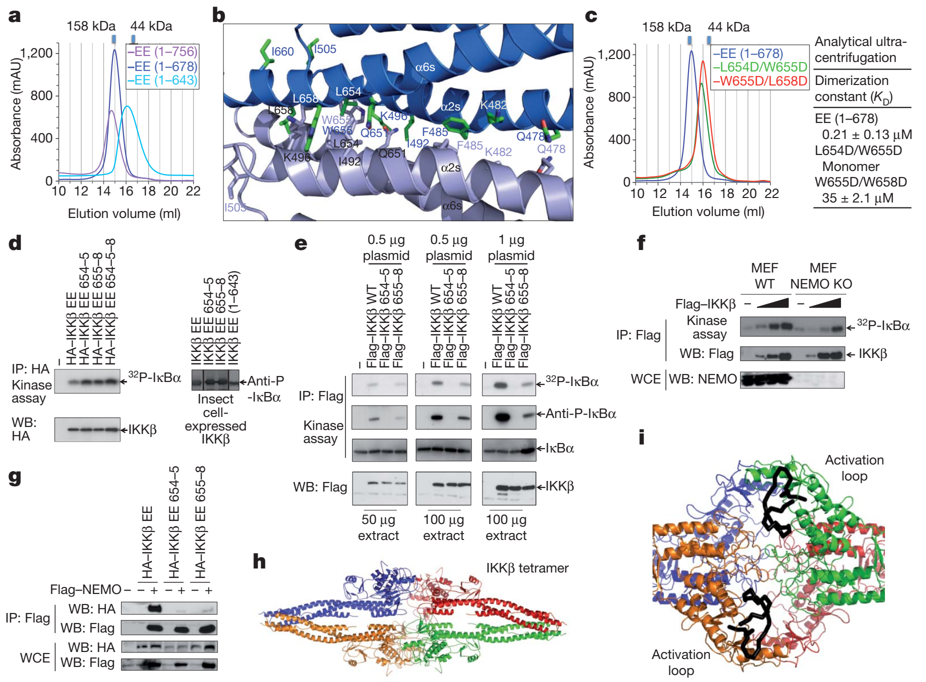


Figure 5 | Dimerization is critical for IKKβ activation but not for its activity. a, Gel filtration profiles of various hIKKβ constructs showing that those containing the KD–ULD–SDD region (1–756, 1–678) are dimeric and that a further truncated construct (1–643) is monomeric. mAU, milliabsorbance unit. **b**, The dimerization interface of xIKKβ. **c**, Structure-based mutations disrupt hIKKβ dimerization as shown by gel filtration and analytical ultracentrifugation. **d**, Kinase activity of HEK293T-cell-transfected or insect-cell-purified hIKKβ EE and dimerization-defective mutants L654D/W655D (654–655), W655D/L658D (655–658) and L654D/W655D/L658D (654–655–658). **e**, Autoactivation of HEK293T-cell-transfected wild-type hIKKβ and its dimerization-defective mutants. **f**, Transfection of hIKKβ into wild-type and NEMO$^{-/-}$ mouse embryonic fibroblasts (MEFs), showing reduced IKKβ activation in the absence of NEMO. **g**, Dimerization mutants of IKKβ showing reduced interaction with NEMO. **h**, Tetramer of xIKKβ in the *P*1 structure. **i**, Close-up view of the tetramer interface, showing that the activation loops of neighbouring protomers (black) face each other.

and the single IKKβ dimer in $I4_122$ shows that IKKβ dimers are conserved (Supplementary Fig. 7). In $I4_122$, the distal part of the SDD is not visible, owing to a lack of crystal packing in this region and dynamic disorder, not degradation (Supplementary Fig. 8).

To confirm the observed interface in IKKβ dimerization, we performed structure-based mutagenesis on residues that bury the most surface area, Leu 654, Trp 655 and Leu 658. These residues are not within the predicted LZ region. Two purified double mutants, L654D/W655D and W655D/L658D, failed to dimerize, as shown by the considerable shift in gel filtration positions (Fig. 5c). Furthermore, equilibrium analytical ultracentrifugation experiments showed that wild-type IKKβ is indeed a dimer whereas one of the IKKβ mutants is a monomer and the other has a 167-fold weaker dimerization affinity (Fig. 5c).

Dimerization in IKKβ activation but not activity

The geometry of the IKKβ dimer with KDs facing away from each other suggests that each IKKβ protomer is independent in its kinase activity. To confirm this, we transfected HEK293T cells with hIKKβ EE mutants, L654D/W655D (654–655), W655D/L658D (655–658) and L654D/W655D/L658D (654–655–658), and found that all mutants had robust kinase activity (Fig. 5d). Dimerization mutants expressed in insect cell and purified showed the same results. In addition, a purified IKKβ construct with truncation into the SDD (1–643) that is monomeric in solution (Fig. 5a) was active in phosphorylating IκBα (Fig. 5d).

Previous studies have shown that IKKβ can undergo *trans*-autophosphorylation and activation on transfection[13,16]. This autophosphorylation and phosphorylation by TAK1 probably both contribute to IKKβ activation on cell stimulation. To determine whether the observed dimerization interface is critical for this activation, we tested dimerization mutations in IKKβ activation on overexpression in HEK293T cells (Fig. 5e). Whereas wild-type IKKβ was robustly activated, the L654D/W655D and W655D/L658D mutants completely and partly lost this activation, respectively, in a manner that correlates with the degree of impairment in dimerization (Fig. 5c, e). Overexpression of IKKβ in NEMO-deficient mouse embryonic fibroblasts resulted in its activation, but to a lesser extent than in wild-type mouse embryonic fibroblasts (Fig. 5f). We found that IKKβ dimerization mutants failed to interact with NEMO efficiently (Fig. 5g), perhaps owing to reduced avidity for oligomeric NEMO. Therefore, although IKKβ kinase activity does not depend on dimerization once its activation loop is phosphorylated, IKKβ activation requires dimerization and is probably enhanced by interaction with NEMO.

Discussion

The IKKβ structure presented here predicts that IKKα, TBK1 and IKK-i all have an overall structural organization that comprises KD, ULD and SDD. Although a ULD was not predicted in IKKα, conservation of buried residues in this region and of the ULD–SDD interface suggests that IKKα also has this domain (Supplementary Fig. 4). The ULD and SDD probably have the same structural role but may have evolved additional, differential functions in the individual kinases. SDD-mediated dimerization may also be conserved. In particular, residues at the observed IKKβ dimerization interface are highly conserved in IKKα (Supplementary Fig. 4), explaining how IKKα and IKKβ can form both homo- and heterodimers[8,13]. Given our structure of IKKβ and the previously determined structures of NEMO fragments[33–37], we may speculate that the full IKK complex is also a dimer, or a dimer of dimers with a molecular mass of around 270 or 540 kD. The apparent 700–900-kD molecular mass of the IKK holoenzyme on gel filtration[5–10] may be due to the elongated nature of NEMO and the complex (Supplementary Fig. 9).

Because the conserved IKKβ dimer structure does not place the KDs close to each other for *trans*-autophosphorylation, we wondered whether higher-order oligomerization, which is probably enhanced by NEMO and its ability to bind ubiquitin[33,34], is responsible for this autoactivation. In both *P*1 and $I4_122$, IKKβ exists as dimers of dimers (Fig. 5h and Supplementary Fig. 10). In particular, active sites of two neighbouring protomers in the tetramer face each other such that an activation loop from one protomer might reach into the active site of the other (Fig. 5i), which may mimic an autophosphorylation state.

The ULD–SDD region of IKKβ directly interacts with the C-terminal portion of IκBα. This interaction may serve several purposes. First, it probably orients IκBα such that its N-terminal cognate phosphorylation sites are presented to the KD active site (Fig. 4g). Without this interaction, IKKβ preferentially phosphorylates the C-terminal PEST motif of IκBα. Second, the interaction seems to enhance IKKβ activity. Third, phosphorylation at the PEST by casein kinase 2 or other kinases may regulate IκBα interaction with IKKβ

and hence affect phosphorylation at the cognate sites. Fourth, the interaction may allow concerted phosphorylation at both Ser 32 and Ser 36 of IκBα without intervening dissociation. In MAP-kinase cascades that involve dual phosphorylations, specific docking interactions occur between the kinases and their substrates[27,38,39]. The β-catenin protein in the Wnt signalling pathway contains the same dual-phosphorylated destruction box motif as IκBα (ref. 40). Consistent with this, β-catenin is brought to the responsible kinase, GSK-3, by means of the adaptor protein axin, allowing both specificity and concerted phosphorylation[41,42]. Therefore, in a general statement that structure serves function, the IKKβ structure seems to fit its function as the IκBα kinase perfectly, being poised to turn on the important NF-κB pathway specifically, efficiently and concertedly in response to cellular physiology.

METHODS SUMMARY

Xenopus laevis IKKβ was expressed in insect cells and purified to homogeneity using nickel affinity chromatography, ion exchange and gel filtration chromatography. We crystallized the $P1$ and $I4_122$ forms at 4 °C in polyethylene glycol 6000 and potassium/sodium phosphate, respectively. The structure was determined by multiwavelength anomalous diffraction of the selenomethionyl protein.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 September 2010; accepted 17 January 2011.
Published online 20 March 2011.

1. Hayden, M. S. & Ghosh, S. Shared principles in NF-κB signaling. *Cell* **132,** 344–362 (2008).
2. Vallabhapurapu, S. & Karin, M. Regulation and function of NF-κB transcription factors in the immune system. *Annu. Rev. Immunol.* **27,** 693–733 (2009).
3. Scheidereit, C. IκB kinase complexes: gateways to NF-κB activation and transcription. *Oncogene* **25,** 6685–6705 (2006).
4. Karin, M. Nuclear factor-κB in cancer development and progression. *Nature* **441,** 431–436 (2006).
5. Chen, Z. J., Parent, L. & Maniatis, T. Site-specific phosphorylation of IκBα by a novel ubiquitination-dependent protein kinase activity. *Cell* **84,** 853–862 (1996).
6. DiDonato, J. A. *et al.* A cytokine-responsive IκB kinase that activates the transcription factor NF-κB. *Nature* **388,** 548–554 (1997).
7. Mercurio, F. *et al.* IKK-1 and IKK-2: cytokine-activated IκB kinases essential for NF-κB activation. *Science* **278,** 860–866 (1997).
8. Woronicz, J. D. *et al.* IκB kinase-β: NF-κB activation and complex formation with IκB kinase-α and NIK. *Science* **278,** 866–870 (1997).
9. Yamaoka, S. *et al.* Complementation cloning of NEMO, a component of the IκB kinase complex essential for NF-κB activation. *Cell* **93,** 1231–1240 (1998).
10. Zandi, E. *et al.* The IκB kinase complex (IKK) contains two kinase subunits, IKKα and IKKβ, necessary for IκB phosphorylation and NF-κB activation. *Cell* **91,** 243–252 (1997).
11. Rothwarf, D. M., Zandi, E., Natoli, G. & Karin, M. IKK-γ is an essential regulatory subunit of the IκB kinase complex. *Nature* **395,** 297–300 (1998).
12. Hacker, H. & Karin, M. Regulation and function of IKK and IKK-related kinases. *Sci. STKE* **2006,** re13 (2006).
13. Zandi, E., Chen, Y. & Karin, M. Direct phosphorylation of IκB by IKKα and IKKβ: discrimination between free and NF-κB-bound substrate. *Science* **281,** 1360–1363 (1998).
14. Sato, S. *et al.* Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nature Immunol.* **6,** 1087–1095 (2005).
15. Liu, H. H., Xie, M., Schneider, M. D. & Chen, Z. J. Essential role of TAK1 in thymocyte development and activation. *Proc. Natl Acad. Sci. USA* **103,** 11677–11682 (2006).
16. Tang, E. D. *et al.* Roles for homotypic interactions and transautophosphorylation in IκB kinase (IKKβ) activation. *J. Biol. Chem.* **278,** 38566–38570 (2003); erratum **278,** 49661 (2003).
17. Knighton, D. R. *et al.* Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253,** 407–414 (1991).
18. Dikic, I., Wakatsuki, S. & Walters, K. J. Ubiquitin-binding domains — from structures to functions. *Nature Rev. Mol. Cell Biol.* **10,** 659–671 (2009).
19. Zheng, J. *et al.* 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr. D* **49,** 362–365 (1993).
20. Bossemeyer, D. *et al.* Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with Mn^{2+} adenylyl imidodiphosphate and inhibitor peptide PKI(5–24). *EMBO J.* **12,** 849–859 (1993).
21. Xu, R. M., Carmel, G., Kuret, J. & Cheng, X. Structural basis for selectivity of the isoquinoline sulfonamide family of protein kinase inhibitors. *Proc. Natl Acad. Sci. USA* **93,** 6308–6313 (1996).
22. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385,** 602–609 (1997).
23. Noble, M. E., Endicott, J. A. & Johnson, L. N. Protein kinase inhibitors: insights into drug design from structure. *Science* **303,** 1800–1805 (2004).
24. Nolen, B., Taylor, S. & Ghosh, G. Regulation of protein kinases; controlling activity through activation segment conformation. *Mol. Cell* **15,** 661–675 (2004).
25. Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* **376,** 313–320 (1995).
26. May, M. J. *et al.* A novel ubiquitin-like domain in IκB kinase β is required for functional activity of the kinase. *J. Biol. Chem.* **279,** 45528–45539 (2004).
27. Goldsmith, E. J. *et al.* Substrate and docking interactions in serine/threonine protein kinases. *Chem. Rev.* **107,** 5065–5081 (2007).
28. Brown, K. *et al.* Structural basis for the interaction of TAK1 kinase with its activating protein TAB1. *J. Mol. Biol.* **354,** 1013–1020 (2005).
29. Ikeda, F. *et al.* Involvement of the ubiquitin-like domain of TBK1/IKK-i kinases in regulation of IFN-inducible genes. *EMBO J.* **26,** 3451–3462 (2007).
30. Kato, T. Jr, Delhase, M., Hoffmann, A. & Karin, M. CK2 is a C-terminal IκB kinase responsible for NF-κB activation during the UV response. *Mol. Cell* **12,** 829–839 (2003).
31. Barroga, C. F., Stevenson, J. K., Schwarz, E. M. & Verma, I. M. Constitutive phosphorylation of I kappa B alpha by casein kinase II. *Proc. Natl Acad. Sci. USA* **92,** 7637–7641 (1995).
32. Shaul, J. D., Farina, A. & Huxford, T. The human IKKβ subunit kinase domain displays CK2-like phosphorylation specificity. *Biochem. Biophys. Res. Commun.* **374,** 592–597 (2008).
33. Lo, Y. C. *et al.* Structural basis for recognition of diubiquitins by NEMO. *Mol. Cell* **33,** 602–615 (2009).
34. Rahighi, S. *et al.* Specific recognition of linear ubiquitin chains by NEMO is important for NF-κB activation. *Cell* **136,** 1098–1109 (2009).
35. Rushe, M. *et al.* Structure of a NEMO/IKK-associating domain reveals architecture of the interaction site. *Structure* **16,** 798–808 (2008).
36. Bagnéris, C. *et al.* Crystal structure of a vFlip-IKKγ complex: insights into viral activation of the IKK signalosome. *Mol. Cell* **30,** 620–631 (2008).
37. Cordier, F. *et al.* Solution structure of NEMO zinc finger and impact of an anhidrotic ectodermal dysplasia with immunodeficiency-related point mutation. *J. Mol. Biol.* **377,** 1419–1432 (2008).
38. Remenyi, A., Good, M. C. & Lim, W. A. Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol.* **16,** 676–685 (2006).
39. Kallunki, T., Deng, T., Hibi, M. & Karin, M. c-Jun can recruit JNK to phosphorylate dimerization partners via specific docking interactions. *Cell* **87,** 929–939 (1996).
40. Wu, G. *et al.* Structure of a β-TrCP1-Skp1-β-catenin complex: destruction motif binding and lysine specificity of the SCF(β-TrCP1) ubiquitin ligase. *Mol. Cell* **11,** 1445–1456 (2003).
41. Ikeda, S. *et al.* Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3β and beta-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. *EMBO J.* **17,** 1371–1384 (1998).
42. Hart, M. J. *et al.* Downregulation of β-catenin by human Axin and its association with the APC tumor suppressor, β-catenin and GSK3β. *Curr. Biol.* **8,** 573–581 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Rajashankar and N. Sukumar for data collection at the NE-CAT of APS, B. Schwer for help with the kinase assay, P. Gaillard for help with the chemistry and G. Ahlsen, L. Shapiro and B. Honig for the ultracentrifugation experiments. This work was supported by the National Institutes of Health (H.W. and M.K.), the American Heart Association (G.X. and Y.-C.L.) and the Cancer Research Institute (Y.-C.L.). M.K. is an American Cancer Society Research Professor.

Author Contributions G.X. cloned, expressed, purified, crystallized and determined the crystal structure of xIKKβ and performed experiments to determine K_m. Y.-C.L. cloned, expressed, purified and crystallized hIKKβ and performed pull-down experiments and kinase assays using phospho-IκBα antibody. Q.L. expressed the hIKKβ mutants in insect cells. G.N. and X.W. performed transfection, immunoprecipitation and kinase assays and M.K. supervised these experiments. H.W. supervised the project. G.X and H.W. made the figures and wrote the manuscript.

Author Information Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 3QA8 and 3QAD. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.W. (haowu@med.cornell.edu).

METHODS

Protein expression and purification. To elucidate the molecular basis of IKKβ function, we expressed IKKβ from a number of species using baculovirus-mediated insect cell expression. The xIKKβ sequence in the NCBI database starts at a Met residue that is equivalent to Met 17 of both the hIKKβ and the mouse IKKβ (mIKKβ) sequence. Translation of the DNA sequence preceding the ATG codon of Met 17 revealed sequences that are almost identical to residues 9–16 of hIKKβ and mIKKβ. These were taken as part of the xIKKβ sequence and residues 1–8 were taken from the corresponding mIKKβ sequence. This reconstructed xIKKβ sequence has the same residue numbering as the hIKKβ sequence until after the SDD.

Various constructs of IKKβ wild type and the phosphomimetic S177E/S181E mutant were designed with an N-terminal polyhistidine tag and a tobacco etch virus protease-cutting site between the tag and the protein. Recombinant IKKβ baculoviruses were made in DH10BAC cells, amplified and used to infect Hi5 insect cells in serum-free media (Invitrogen). The cells were cultured in suspension and harvested 48 h after infection. The recombinant proteins were purified by nickel affinity chromatography, anion exchange and gel filtration chromatography. For crystallization, the polyhistidine tag was cleaved by the tobacco etch virus protease during protein purification.

All IκBα proteins were expressed in *Escherichia coli* using pET28a, pGEX4T3 and pET-SUMO vectors and purified by their respective affinity tags. For kinase assays, the SUMO tag was cleaved from IκBα proteins. His–ULD and His–ULD–SDD of hIKKβ were also expressed in *E. coli* using the pET28a vector.

Crystallization and data collection. Unlike many protein kinases, the IKKβ KD cannot be recombinantly expressed as a well-behaved biochemical entity for structural studies. In addition, after mapping a compact region by limited proteolysis, IKKβ was still refractory to crystallization, both alone and in the presence of various ATP analogues. To overcome this obstacle, we used several IKKβ inhibitors, including Cmpd1 and Cmpd2 (Supplementary Fig. 1), which were identified against the S177E/S181E (EE) mutant, in co-crystallization. A hIKKβ EE construct (1–678) lacking only the C-terminal NBD did crystallize; however, these crystals only diffracted to a resolution of ~7.5 Å. Searching IKKβ orthologues that may give better crystals led to success in crystallizing the analogous region of xIKKβ EE (4–675; Fig. 1a).

The xIKKβ (S177E/S181E) protein containing residues 4–675 was concentrated by ultrafiltration (Amicon) to about 15 mg ml^{-1} in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 10 mM dithiothreitol (DTT). It was mixed with an inhibitor compound in a 1:2 molar ratio before crystallization. Cmpd1 is 4-((4-(4-(chlorophenyl)pyrimidin-2-yl)amino)phenyl)(4-(2-hydroxyethyl)piperazin-1-yl)methanone and Cmpd2 is 1-(4-(4-((4-(4-(pyridin-4-ylsulfonyl)phenyl)-pyrimidin-2-yl)amino)benzoyl)piperazin-1-yl)ethanone. The *P*1 crystals were grown using hanging-drop vapour diffusion at 4 °C by mixing equal volumes of the purified protein and the crystallization condition of 100 mM *N*-(2-acetamido) iminodiacetic acid at pH 6.5, 10% (w/v) polyethylene glycol 6000, 50 mM Li_2SO_4, 300 mM NaCl and 10 mM DTT. The $I4_122$ crystals were grown at 4 °C with well solution containing 1.8 M K/Na phosphate at pH 5.6 and 10 mM DTT. For data collection, all crystals were flash frozen in the respective crystallization conditions supplemented with 25% (v/v) ethylene glycol. Diffraction data were collected at the 24ID-C beam line of the Advanced Photon Source. Multiwavelength anomalous diffraction (MAD) data on heavy-atom derivative crystals or selenomethionyl crystals were collected near the respective absorption edges. All diffraction data were processed using the HKL2000 suite[43] and their statistics are shown in Supplementary Table 1 and Supplementary Table 2.

Structure determination, refinement and analysis. The initial xIKKβ crystals grew in the *P*1 space group in the presence of the inhibitor Cmpd1 or Cmpd2 and diffracted to a resolution of 3.6 Å. Selenomethionyl crystals were obtained, but we failed to locate the large number of expected selenium sites. Among the extensive heavy-atom searches, an ytterbium-derivative was obtained, with eight well-defined sites, which probably correspond to eight IKKβ molecules in the asymmetric unit. However, the electron density map calculated from a three-wavelength ytterbium-anomalous diffraction data set was insufficient for tracing, and phase combination with the selenomethionyl data set could not be performed, owing to non-isomorphism.

The structure determination was eventually successful in the alternative crystal form, $I4_122$, which contains one molecule of IKKβ in complex with Cmpd1 and diffracted to a resolution of 4.0 Å, using MAD of the selenomethionyl crystals (Supplementary Tables 1 and 2 and Supplementary Fig. 2). Twelve selenium sites were determined using the program SHELXD[44] and refined with the program MLPHARE in the CCP4 suite[45]. MAD phases were calculated at a resolution of 4.0 Å with data from $I4_122$ crystals using the program SHARP[46]. A few cycles of model building and refinement were carried out with the program O[47] and REFMAC with TLS parameterization[48]. The $I4_122$ crystals contain one monomer in the asymmetric unit and 80% solvent when calculated with the entire IKKβ construct and 84% solvent when considering only the ordered part of the structure. The inhibitor Cmpd1 has density in the MAD experimental map and the $F_o - F_c$ omit map. The Dundee PRODRG2 Server was used to generate topology and restraint files of the compound for refinement. The refined model contains residues 16–236, 243–286, 290–376, 384–394, 401–475 and 528–637 and Cmpd1.

The structure of the *P*1 form was determined by molecular replacement using the refined model of the $I4_122$ crystal form as the search model, in which eight molecules were located. Selenium sites of the single-wavelength anomalous diffraction data set of a selenomethionyl crystal in the *P*1 space group was calculated by difference Fourier analysis and used to verify residue registration in the *P*1 structure. Refinement in the *P*1 structure was conducted at a resolution of 3.6 Å and incorporated tight non-crystallographic symmetry restraints (root mean squared deviation of 0.02 Å in atom positions). After several rounds of refinement at a resolution of 3.6 Å, new electron densities appeared in the *P*1 crystal form to complete the model building. Although Cmpd2 was in the crystallization condition, it did not have clear density and was not included in the refinement. The refined model of the *P*1 crystal form contains four IKKβ dimers in the asymmetric unit. Three of the dimers encompass residues 16–236, 243–286, 290–376, 384–394, 401–551 and 559–666. One dimer contains the same residues as the structure in $I4_122$. The structures were analysed using the CCP4 suite[45] and the Dali server[49], and the figures were made using PYMOL[50].

GST pull-down. The tagged proteins were first purified with glutathione or Ni-NTA beads and their expression levels were assessed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE). Beads containing estimated equivalent quantities of the tagged proteins were mixed with the cell lysates or the purified versions of the interaction partners. The mixtures were incubated at room temperature (20 °C) for 1 h with rotation. After centrifugation, the supernatants were removed. The beads were then washed twice, eluted and subjected to SDS–PAGE analysis. All pull-down experiments were repeated two to four times with consistency.

Kinase assays using anti-phospho-IκBα antibody. The hIKKβ proteins (0.1 μg μl^{-1}) were incubated with recombinant IκBα (1 μg μl^{-1}) in 50 mM Tris-HCl at pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$ and 2 mM DTT for 30 min at 30 °C. SDS–PAGE sample buffer was used to terminate the reactions. The products were separated on 15% SDS–PAGE and transferred to PVDF membranes. Anti-phospho-IκBα antibody (Cell Signaling Technology) was used to detect phospho-IκBα.

Determination of K_m of hIKKβ for IκBα (1–54) and full-length IκBα. To derive the K_m of hIKKβ for full-length IκBα, kinase assays were performed at substrate concentrations of 2.8, 5.3, 10.6, 21.3, 42.5 and 85 μM. Similarly, to derive the K_m of hIKKβ for IκBα (1–54), kinase assays were performed at substrate concentrations of 5.3, 10.6, 21.3, 42.5, 85 and 170 μM. A time course of the kinase reactions was first performed to select a hIKKβ amount and a time point within which the reactions are linear with time. The final selected reactions contain 10 ng baculovirus-expressed hIKKβ, 100 mM cold ATP and 1 μl [γ-^{32}P]ATP (3,000 Ci mmol^{-1}, 1 mCi per 100 μl) in 25 μl of reaction buffer containing 50 mM Tris-HCl at pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$ and 2 mM DTT. The phosphotransfer reaction was allowed to proceed for 10 min at 30 °C and quenched with SDS–PAGE sample buffer. The products were separated on 15% SDS–PAGE and subjected to autoradiography. The relative amounts of phosphorylated IκBα were quantified using IMAGEJ, plotted against total IκBα concentrations and fitted using nonlinear regression to the Michaelis–Menten equation to obtain K_m using SIGMAPLOT.

Transfection, immunoprecipitation and kinase assay. The constructs Flag–hIKKβ EE and its truncation mutants; HA–hIKKβ EE and its dimerization mutants L654D/W655D, W655D/L658D and L654D/W655D/L658D; and Flag–hIKKβ and its dimerization mutants L654D/W655D and W655D/L658D were generated in the vector pcDNA3 by conventional PCR. All IKKβ constructs were transfected in HEK293T cells with Lipofectamine 2000 (Invitrogen). After 24 h, cell extracts were immunoprecipitated with anti-Flag antibodies bound to agarose beads (M2, Sigma) or anti-HA bound to agarose beads (Sigma). IKKβ kinase assays were essentially done as previously described[6,13]. Briefly, immunoprecipitates were incubated with 2 μM full-length IκBα (1–317) in 20 mM HEPES at pH 7.5, 10 mM $MgCl_2$, 20 mM β-glycerophosphate, 10 mM PNPP, 50 mM Na_3VO_4, 1 mM DTT, 20 mM ATP, and 1–10-mCi [γ-^{32}P]ATP at 30 °C for 30 min, and subjected to SDS–PAGE and autoradiography. Immunoblotting was performed using anti-Flag (Sigma), anti-HA (Sigma) or anti-IKKβ antibodies (Upstate, 05-535).

Equilibrium analytical ultracentrifugation measurements. Experiments were performed in a Beckman XL-A/I analytical ultracentrifuge (Beckman-Coulter), using six-cell centre pieces with straight walls, a 12-mm path length and sapphire windows. Samples were kept and diluted in 50 mM Tris-HCl at pH 8.0 and 300 mM NaCl. Samples from wild-type protein were diluted to 6.9, 4.5 and 2.4 μM, mutant L654D/W655D to 7.4, 4.8 and 2.6 μM and mutant W655D/L658D to 4.9, 3.2 and 1.7 μM for channels A, B and C, respectively. Dilution buffer was used as blank. All samples were run at 4 °C at 9,000 r.p.m. (5,900*g*; held for 20 h then scanned four times at 1-h intervals), 11,000 r.p.m. (8,800*g*; held for 10 h then scanned four times at 1-h intervals), 14,000 r.p.m. (14,300*g*; held for

10 h then scanned four times at 1-h intervals) and 17,000 r.p.m. (21,000*g*; held for 10 h then scanned four times at 1-h intervals). Detection was by ultraviolet absorption at 280 nm. Solvent density and the protein partial specific volume at each temperature were determined using the program SEDNTERP (Alliance Protein Laboratories). For calculation of K_D and the apparent molecular weight, all useful data were used in a global fit, using the program HETEROANALYSIS, obtained from University of Connecticut (http://www.biotech.uconn.edu/auf).

43. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
44. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58,** 1772–1779 (2002).
45. Collaborative Computational Project, Number 4. The *CCP4* suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
46. Bricogne, G. *et al.* Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59,** 2023–2030 (2003).
47. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building models in electron density maps and the location of errors in those models. *Acta Crystallogr. A* **47,** 110–119 (1991).
48. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374,** 300–321 (2003).
49. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20,** 478–480 (1995).
50. DeLano, W. L. *PyMOL Molecular Viewer* ⟨http://www.pymol.org⟩ (2002).

doi:10.1038/nature09928

The auroral footprint of Enceladus on Saturn

Wayne R. Pryor[1,2]*, Abigail M. Rymer[3]*, Donald G. Mitchell[3], Thomas W. Hill[4], David T. Young[5], Joachim Saur[6], Geraint H. Jones[7,8], Sven Jacobsen[6], Stan W. H. Cowley[9], Barry H. Mauk[3], Andrew J. Coates[7], Jacques Gustin[10], Denis Grodent[10], Jean-Claude Gérard[10], Laurent Lamy[11], Jonathan D. Nichols[9], Stamatios M. Krimigis[3,12], Larry W. Esposito[13], Michele K. Dougherty[14], Alain J. Jouchoux[13], A. Ian F. Stewart[13], William E. McClintock[13], Gregory M. Holsclaw[13], Joseph M. Ajello[15], Joshua E. Colwell[16], Amanda R. Hendrix[15], Frank J. Crary[5], John T. Clarke[17] & Xiaoyan Zhou[15]

Although there are substantial differences between the magnetospheres of Jupiter and Saturn, it has been suggested that cryovolcanic activity at Enceladus[1–9] could lead to electrodynamic coupling between Enceladus and Saturn like that which links Jupiter with Io, Europa and Ganymede. Powerful field-aligned electron beams associated with the Io–Jupiter coupling, for example, create an auroral footprint in Jupiter's ionosphere[10,11]. Auroral ultraviolet emission associated with Enceladus–Saturn coupling is anticipated to be just a few tenths of a kilorayleigh (ref. 12), about an order of magnitude dimmer than Io's footprint and below the observable threshold, consistent with its non-detection[13]. Here we report the detection of magnetic-field-aligned ion and electron beams (offset several moon radii downstream from Enceladus) with sufficient power to stimulate detectable aurora, and the subsequent discovery of Enceladus-associated aurora in a few per cent of the scans of the moon's footprint. The footprint varies in emission magnitude more than can plausibly be explained by changes in magnetospheric parameters—and as such is probably indicative of variable plume activity.

There have been 12 close Cassini encounters with Enceladus in the six years since the spacecraft arrived at Saturn. During a fly-by on 11 August 2008, the spacecraft passed within 55 km of the moon at 21:06 UT. Cassini approached Enceladus from downstream (with respect to the background plasma flow) while moving north–south (Supplementary Fig. 1). Just before closest approach, a spacecraft roll brought two plasma sensors into the optimum orientation for measuring along Saturn's (approximately dipolar) magnetic field lines. At this time, powerful ion and electron beams were observed propagating from Saturn's northern hemisphere (Fig. 1). Neither sensor was accessible to particles originating from Saturn's southern ionosphere. Beams were observed from 3.6 to at least 23.3 R_E (radius of Enceladus $R_E = 252$ km) downstream (positive X in Fig. 1) from Enceladus. At 21:05 UT, ~1 min before closest approach, with Cassini still 3.6 R_E downstream of the moon, the flow of magnetic-field-aligned ions and electrons abruptly ceased. (The final burst of low energy electron flux observed after closest approach at ~21:07 in Fig. 1b is actually the tail of a non-field-aligned distribution and is produced by a different process to that which produces the beams.)

At approximately 20:59 and 21:02 UT, the magnetic-field-aligned electrons flicker in energy between peaks near 10 eV and 1 keV; bimodal electron populations are observed for about 1 min either side of these transitions (Fig. 1b). These changes in the characteristic energy of the field-aligned electron flux, while not currently well understood, are associated with changes in the magnetic field perturbation (Fig. 1c), suggesting an actual change in the total field-aligned current density. At Jupiter, variations in auroral radio emission[14] and a 'string-of-pearls' ultraviolet aurora associated with the Io footprint[15] have been interpreted as being due to multiple reflections of a standing Alfvén wave current system driven by Io. It is possible that the flickering in energy of the beams observed downstream of Enceladus is the equatorial signature of a standing wave pattern like that observed at the Io footprint. It has been suggested that the locations of beams observed near Io are controlled by the product of the Alfvén wave travel time towards Jupiter and the plasma convection speed past the moon[16]. If this value, in units of the moon's radius, is larger at Saturn, then the beams are expected to be further downstream than at Io.

We estimate the wave travel time to be of the order of 150 s (using the electron density derived from Cassini data[17] and assuming a dipole field). Assuming an average plasma velocity of ~20% of full co-rotation between Enceladus and the onset of the beams, we find a downstream shift of the beams of 3.6 R_E. This is consistent with the observed distance downstream from the moon where the beams begin. However, given the spatial[1] and (likely) temporal[18,19] variability of the Enceladus vents, filamentary current structures associated with local variable mass loading might contribute to the variability of the observations. Assuming the field-aligned electrons are incident on Saturn's ionosphere, the observed flux excites hydrogen molecules at Enceladus' footpoint, producing ultraviolet emission between 3 ± 0.2 and 12 ± 3.0 kR. That is above the measurement threshold of the Cassini UltraViolet Imaging Spectrograph (UVIS)[20].

Two weeks later, on 26 August 2008, the UVIS recorded three successive polar views (two of which are shown here as Fig. 2) that show an unambiguous auroral footprint (boxed area at top left of Fig. 2a and b). UVIS spectra of the footprint look similar to the simultaneously measured emissions from the brighter main auroral oval seen near 75° latitude in Fig. 2. Both compare well with an H_2 electron-impact laboratory spectrum and are thus consistent with emissions due to electrons precipitating on atomic and molecular hydrogen at an emission altitude ~1,100 km above the 1 bar level in the atmosphere[21]. Using this altitude and a quantitative Saturn magnetic field model[22], we calculate that the northern Enceladus footprint should occur at a latitude of 64.5° N for nominal magnetospheric conditions. (The southern footprint would occur at 61.7° S because Saturn's magnetic dipole, although spin-aligned within observational uncertainties, is displaced about 0.04 R_S (Saturn radii) northward from Saturn's geometric centre[22].) The modelled footprint latitude is not

[1]Science Department, Central Arizona College, Coolidge, Arizona 85128 USA. [2]Space Environment Technologies, Pacific Palisades, California 90272, USA. [3]Applied Physics Laboratory, Johns Hopkins University, Laurel, Maryland 20723, USA. [4]Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA. [5]Space Science and Engineering Division, Southwest Research Institute, San Antonio, Texas 78238, USA. [6]Institut für Geophysik und Meteorologie, Universität zu Köln, Cologne, D-50923, Germany. [7]Mullard Space Science Laboratory, Department of Space and Climate Physics, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK. [8]The Centre for Planetary Sciences at University College London/Birkbeck, London WC1E 6BT, UK. [9]Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK. [10]Laboratoire de Physique Atmosphérique et Planétaire, Département d'Astrophysique, Géophysique et Océanographie, Université de Liège, Liège, B-4000, Belgium. [11]Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris, Centre National de la Recherche Scientifique, Université Pierre et Marie Curie, Université Paris Diderot, 92195 Meudon, France. [12]Academy of Athens, Soranou Efesiou 4, 115 27, Athens, Greece. [13]Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, Colorado 80303, USA. [14]Space and Atmospheric Physics, The Blackett Laboratory, Imperial College, London SW7 2AZ, UK. [15]Jet Propulsion Laboratory, Pasadena, California 91109, USA. [16]Department of Physics, University of Central Florida, Orlando, Florida 32816, USA. [17]Astronomy Department, Boston University, Boston 02215, USA.
*These authors contributed equally to this work.

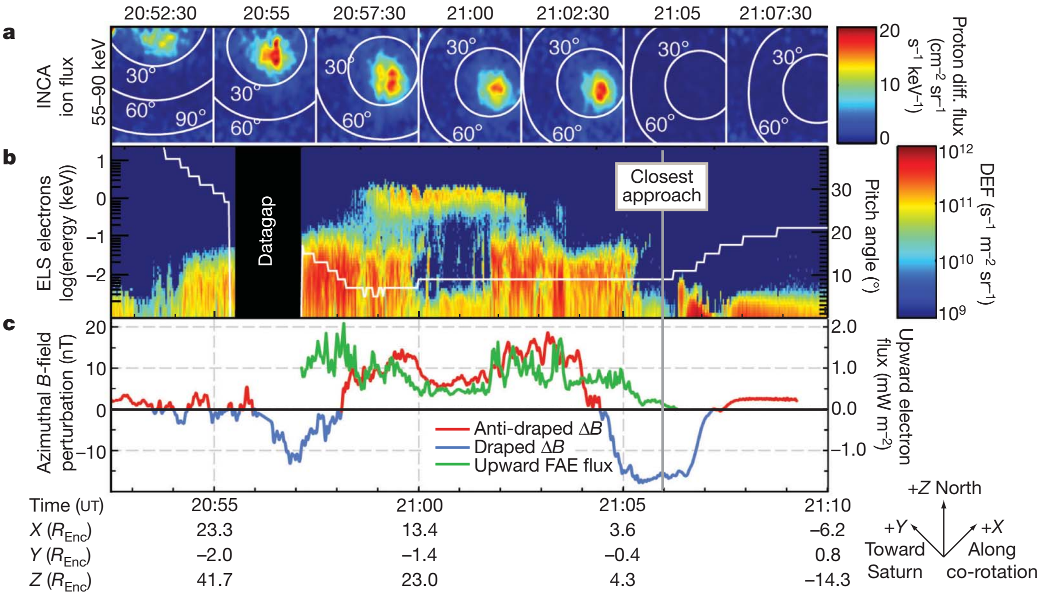


Figure 1 | Cassini particle and field observations on 11 August 2008. **a**, Protons (55–90 keV) observed by the Cassini Ion and Neutral Camera (INCA)[26]. Colours indicate proton differential flux. Contours 30° and 60° from the magnetic field direction are overplotted in white. **b**, Electrons (1 eV to 22 keV) measured by the most field-aligned detectors of the CAPS Electron Spectrometer (ELS)[27]. Electron differential energy flux (DEF) is indicated by the colour bar. Field-aligned electrons are observed when the instrument measured within ∼20° of the magnetic field line, as indicated by the white line (pitch angle). The blacked out regions are data gaps. **c**, Azimuthal perturbation (ΔB) in the magnetic field during this interval, calculated by subtracting a model background field from the total field measured by the Cassini magnetometer (MAG)[28,29]. Positive ΔB (red) is in the direction of co-rotation; the signature expected from simple field line draping around an obstacle is characterized by negative ΔB (blue) above the equator. Positive ΔB (red) indicates an anti-draped perturbation in the super-co-rotational sense. Overplotted in green is the total upward field-aligned electron (FAE) energy flux derived by numerical integration of the electron data in **b** (see Supplementary Information for additional discussion and error analysis). Calculations of electron energy loss in H_2 atmospheres indicate that 1 mW m^{-2} of particle energy input produces ∼10 kR of auroral ultraviolet emission[30]. The observed electron energy flux is therefore expected to produce a ultraviolet emission brightness between 2.8 ± 0.2 and 11.9 ± 3.0 kR.

very sensitive to auroral altitude; it would shift only 0.04° equatorward if the assumed auroral altitude were increased to 1,200 km. It is also not very sensitive to the size of the magnetospheric cavity, varying by only ∼0.1° over the whole range of sizes observed during the 6-year Cassini mission.

The location of the observed northern footprint is consistent with the expected location. The brightness centroid of the first spot (Fig. 2a) was at latitude 64.1° ± 0.4° and longitude 286.0° ± 0.5°, thus about 1.7° downstream of the sub-Enceladus longitude of 287.7°. Here we have set errors equal to the pixel size. The brightness centroid of the

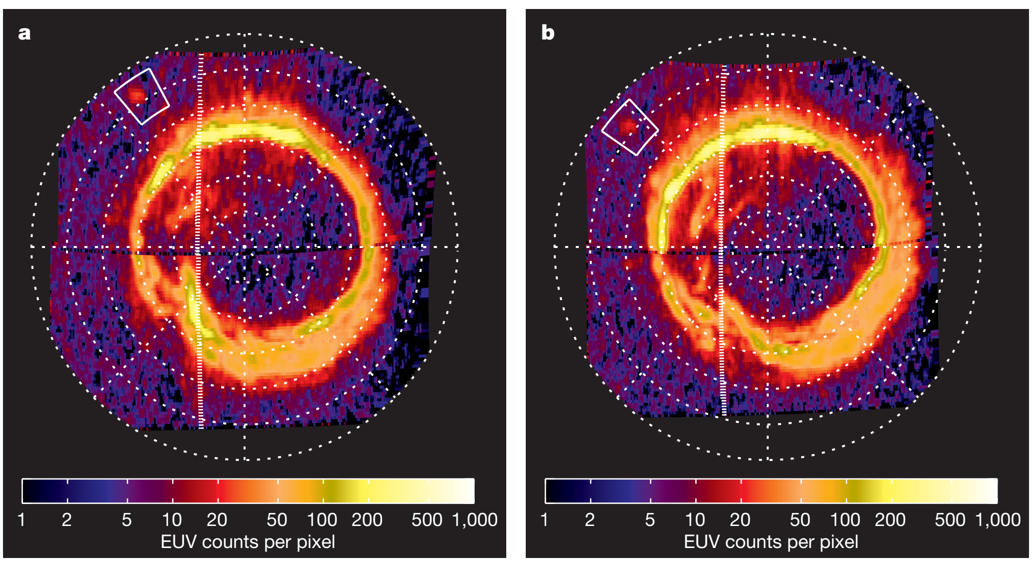


Figure 2 | Cassini images of Saturn's northern aurora, including the Enceladus auroral footprint. **a**, **b**, Successive UVIS EUV polar-projected images of Saturn's north polar region from 26 August 2008 (day of year 239); 02:16–03:28 UT (**a**) and 03:38–04:50 UT (**b**). Images were formed by slowly slewing the spacecraft and its long-slit ultraviolet spectrometer. During this interval, Cassini moved from sub-spacecraft latitudes of 74° N to 65° N, and from 8.1 to 6.0 R_S (Saturn radius $R_S \approx$ 60,300 km) from Saturn's centre. Each image represents two spacecraft slews across the planet. The colour bar shows EUV emission per pixel. The white boxes are centred on 64.5° N and the sub-Enceladus longitude, cover 4° in latitude and 10° in longitude, and enclose the predicted magnetic mapping of the satellite Enceladus to Saturn's dayside atmosphere. Satellite footprint emission is visible in both boxes. The north pole is at the centre; the latitude circles are 5° apart, and the hashed white line indicates the day/night terminator. The Sun is to the left.

second spot was at latitude 64.3° ± 0.3° and longitude 316.8° ± 0.4°, about 0.8° downstream of the sub-Enceladus longitude of 317.6°. These values are very close to the anticipated Enceladus footprint latitude of 64.5°. This agreement confirms both the identification with Enceladus and the accuracy of the magnetic field model, a 'ground-truth' observation which proved vital in confirming the detailed magnetic field configuration of Jupiter[23] and is equally valuable at Saturn. The predicted southern footprint has not yet been detected (28 of the 310 non-detections were of the southern ionosphere); the southern footprint may be dimmer than its northern counterpart, as is the case for the main aurora[24].

The 504 km diameter of Enceladus maps along Saturn's magnetic field lines to a quasi-elliptical spot ~52 × 29 km in Saturn's atmosphere, for average magnetospheric conditions. This spot would be spatially unresolved by UVIS. For a steady fixed source, the UVIS observations suggest emission connected to an Enceladus interaction region at the equator extending as far as 20 R_E downstream with a radial extent between 0 and 20 R_E, consistent with the location of the beams observed at the equator.

The slewing UVIS slit passed over and recorded the Enceladus-related spot on 26 August 2008 at 3:00, at 4:20 and at 8:02 UT. The spot dims as it moves from near dawn towards noon (8:20 to 9:12 to 11:54 local time). The total combined (extreme ultraviolet (EUV) plus far ultraviolet (FUV)) spot brightnesses in the three UVIS images were 1,550 ± 340 R, 1,130 ± 200 R and 450 ± 290 R. These should be considered lower limits, assuming the spatial pixel is uniformly filled by signal, because the true emission region size is not known (see Supplementary Information for more details). Thus (even when visible) the Enceladus auroral footprint varies in brightness by a factor of about 3. At Jupiter, footprint emission variability is principally caused by the rocking of the magnetospheric plasma sheet, as the magnetic dipole moment is inclined with respect to the spin axis. At Saturn, there is no substantial rocking of the plasma sheet at the location of Enceladus, but we still see these large brightness variations. This variation could, in principle, reflect variations of plume activity[18,19], of ionization rates (owing to varying background plasma conditions), or of magnetospheric size (when the magnetosphere is compressed, auroral emissions are generally enhanced). The last two factors do not typically exhibit order-of-magnitude variations[17,25]. The most likely cause for the observed large-scale variability, therefore, is time-variable cryovolcanism from Enceladus' south polar vents, suggesting that plume activity was particularly high during August 2008. Thus, systematic monitoring of Enceladus' ultraviolet auroral footprint might provide evidence of plume variability, which is an important open issue.

Received 13 July 2010; accepted 10 February 2011.

1. Porco, C. C. *et al.* Cassini observes the active south pole of Enceladus. *Science* **311,** 1393–1401 (2006).
2. Spencer, J. R. *et al.* Cassini encounters Enceladus: background and the discovery of a south polar hot spot. *Science* **311,** 1401–1405 (2006).
3. Dougherty, M. K. *et al.* Identification of a dynamic atmosphere at Enceladus with the Cassini magnetometer. *Science* **311,** 1406–1409 (2006).
4. Tokar, R. L. *et al.* The interaction of the atmosphere of Enceladus with Saturn's plasma. *Science* **311,** 1409–1412 (2006).
5. Jones, G. H. *et al.* Enceladus' varying imprint on the magnetosphere of Saturn. *Science* **311,** 1412–1415 (2006).
6. Spahn, F. *et al.* Cassini dust measurements at Enceladus and implications for the origin of the E ring. *Science* **311,** 1416–1418 (2006).
7. Waite, J. H. *et al.* Cassini ion and neutral mass spectrometer: Enceladus plume composition and structure. *Science* **311,** 1419–1422 (2006).
8. Hansen, C. J. *et al.* Enceladus' water vapor plume. *Science* **311,** 1422–1425 (2006).
9. Brown, R. H. *et al.* Composition and physical properties of Enceladus' surface. *Science* **311,** 1425–1428 (2006).
10. Connerney, J. E. P., Baron, R., Satoh, T. & Owen, T. Images of excited H_3^+ at the foot of the Io flux tube in Jupiter's atmosphere. *Science* **262,** 1035–1038 (1993).
11. Clarke, J. T. *et al.* Ultraviolet emissions from the magnetic footprints of Io, Ganymede, and Europa on Jupiter. *Nature* **415,** 997–1000 (2002).
12. Pontius, D. H. Jr & Hill, T. W. Enceladus: a significant plasma source for Saturn's magnetosphere. *J. Geophys. Res.* **111,** A09214, doi:10.1029/2006JA011674 (2006).
13. Wannawichian, S., Clarke, J. T. & Pontius, D. H. Jr. Interaction evidence between Enceladus' atmosphere and Saturn's magnetosphere. *J. Geophys. Res.* **113,** A07217, doi:10.1029/2007JA012899 (2008).
14. Gurnett, D. A. & Goertz, C. K. Multiple Alfven wave reflections excited by Io, origin of the Jovian decametric arcs. *J. Geophys. Res.* **86** (A2), 717–722 (1981).
15. Bonfond, B. *et al.* UV Io footprint leading spot: a key feature for understanding the UV Io footprint multiplicity? *Geophys. Res. Lett.* **35,** L05107, doi:10.1029/2007GL032418 (2008).
16. Jacobsen, S. J. *et al.* Location and spatial shape of electron beams in Io's wake. *J. Geophys. Res.* **115,** A04205, doi:10.1029/2009JA014753 (2010).
17. Persoon, A. M., Gurnett, D. A., Kurth, W. S. & Groene, J. B. A simple scale height model of the electron density in Saturn's plasma disk. *Geophys. Res. Lett.* **33,** L18106, doi:10.1029/2006GL027090 (2006).
18. Saur, J. *et al.* Evidence for temporal variability of Enceladus' gas jets: modeling of Cassini observations. *Geophys. Res. Lett.* **35,** L20105, doi:10.1029/2008GL035811 (2008).
19. Smith, H. T. *et al.* Enceladus plume variability and the neutral gas densities in Saturn's magnetosphere. *J. Geophys. Res.* **115,** A10252, doi:10.1029/2009JA015184 (2010).
20. Esposito, L. W. *et al.* the Cassini ultraviolet imaging spectrograph investigation. *Space Sci. Rev.* **115,** 299–361 (2004).
21. Gérard, J.-C. *et al.* Altitude of Saturn's aurora and its implications for the characteristic energy of precipitated electrons. *Geophys. Res. Lett.* **36,** L02202, doi:10.1029/2008GL036554 (2009).
22. Burton, M. E., Dougherty, M. K. & Russell, C. T. Model of Saturn's internal planetary magnetic field based on Cassini observations. *Planet. Space Sci.* **57,** 1706–1713 (2009).
23. Connerney, J. E. P., Acuñna, M. H., Ness, N. F. & Satoh, T. New models of Jupiter's magnetic field constrained by the Io flux tube footprint. *J. Geophys. Res.* **103,** 11929–11939 (1998).
24. Nichols, J. D. *et al.* Saturn's equinoctial auroras. *Geophys. Res. Lett.* **36,** L24102, doi:10.1029/2009GL041491 (2009).
25. Kanani, S. J. *et al.* A new form of Saturn's magnetopause using a dynamic pressure balance model, based on *in situ*, multi-instrument Cassini measurements. *J. Geophys. Res.* **115,** A06207, doi:10.1029/2009JA014262 (2010).
26. Krimigis, S. M. *et al.* Magnetospheric imaging instrument (MIMI) on the Cassini mission to Saturn/Titan. *Space Sci. Rev.* **114,** 233–329 (2004).
27. Young, D. T. *et al.* Cassini plasma spectrometer investigation. *Space Sci. Rev.* **114,** 1–112 (2004).
28. Dougherty, M. K. *et al.* The Cassini magnetic field investigation. *Space Sci. Rev.* **114,** 331–383 (2004).
29. Jia, Y.-D. *et al.* Time varying magnetospheric environment near Enceladus as seen by the Cassini magnetometer. *Geophys. Res. Lett.* **37,** L09203, doi:10.1029/2010GL042948 (2010).
30. Waite, J. Jr *et al.* Electron precipitation and related aeronomy of the Jovian thermosphere and ionosphere. *J. Geophys. Res.* **88,** 6143–6163 (1983).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge support from the NASA/ESA Cassini Project and NASA's Cassini Data Analysis Program.

Author Contributions A.M.R. and W.R.P. discovered the electron beams and the auroral footprint, respectively, and wrote most of the paper. D.G.M. discovered the ion beams and contributed to the text and interpretation. T.W.H. contributed extensively to the text and interpretation. D.T.Y. is CAPS PI and contributed extensively to the text and interpretation. J.S., G.H.J., S.J., B.H.M. and A.J.C. advised on the interpretation of the *in situ* data. S.W.H.C. performed the field line mapping and provided advice on the paper. J.G., D.G., J.-C.G., L.L. and J.D.N. advised on the interpretation of the UVIS data. S.M.K. is the MIMI PI and oversaw the ion data. M.K.D. is the MAG PI and oversaw the magnetometer data. L.W.E. is the UVIS PI and oversaw the UVIS data. A.J.J. and F.J.C. designed the auroral observation campaign. A.I.F.S., W.E.M., J.M.A., J.E.C. and A.R.H. helped to process the UVIS data. J.T.C. provided advice on the HST observations. X.Z. contributed to auroral discussions related to comparisons with terrestrial auroral processes.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.M.R. (abigail.rymer@jhuapl.edu).

LETTER

doi:10.1038/nature09963

Optically healable supramolecular polymers

Mark Burnworth[1], Liming Tang[1], Justin R. Kumpfer[1], Andrew J. Duncan[2], Frederick L. Beyer[2], Gina L. Fiore[3], Stuart J. Rowan[1] & Christoph Weder[1,3]

Polymers with the ability to repair themselves after sustaining damage could extend the lifetimes of materials used in many applications[1]. Most approaches to healable materials require heating the damaged area[2–4]. Here we present metallosupramolecular polymers that can be mended through exposure to light. They consist of telechelic, rubbery, low-molecular-mass polymers with ligand end groups that are non-covalently linked through metal-ion binding. On exposure to ultraviolet light, the metal–ligand motifs are electronically excited and the absorbed energy is converted into heat. This causes temporary disengagement of the metal–ligand motifs and a concomitant reversible decrease in the polymers' molecular mass and viscosity[5], thereby allowing quick and efficient defect healing. Light can be applied locally to a damage site, so objects can in principle be healed under load. We anticipate that this approach to healable materials, based on supramolecular polymers and a light–heat conversion step, can be applied to a wide range of supramolecular materials that use different chemistries.

The healing of cracks in amorphous polymers by heating above the glass transition temperature (T_g) involves surface rearrangement and approach of polymer chains, followed by wetting, diffusion and re-entanglement of the chains[6]. Because the rates of the final two steps are inversely proportional to the molecular mass, healing is generally slow and inefficient. This problem can be overcome by exploiting thermally reversible, covalent bonds[7,8] or non-covalent supramolecular motifs[5,9,10] that allow the reaction equilibrium to be temporarily shifted to lower-molecular-mass species[11] on exposure to heat. This reduces the viscosity of the material, such that defects can be mended, before the equilibrium is shifted back and the polymer is reformed. Supramolecular polymers that phase separate into physically cross-linked networks (Fig. 1a) should be especially well suited for this purpose, because such morphologies generally bestow the material with high toughness. The supramolecular motifs can disengage in the solid state on exposure to heat or a competitive binding agent[12,13], causing disassembly into small molecules[14] and viscosity reductions. Reporting a series of supramolecular materials formed by metal–ligand interactions, we demonstrate here that this architecture is an excellent basis for elastomeric materials in which defects can be efficiently repaired. We show that the use of light[15] as a stimulus for the dissociation of supramolecular motifs has distinct advantages over thermally healable systems, including the possibility of exclusively exposing and healing the damaged region.

The new polymers are based on a macromonomer comprising a rubbery, amorphous poly(ethylene-*co*-butylene) core with 2,6-bis(1′-methylbenzimidazolyl)pyridine (Mebip) ligands at the termini (Fig. 1b, **3**). This design was based on the assumption that the hydrophobic core and the polar metal–ligand motif would phase separate[16]. Metal–Mebip complexes were previously used to self-assemble polymeric materials[17–21], and their optical properties seemed appropriate for optical healing[22] (see below). To probe how the metal ion affects the materials' properties, **3** was self-assembled with $Zn(NTf_2)_2$ or

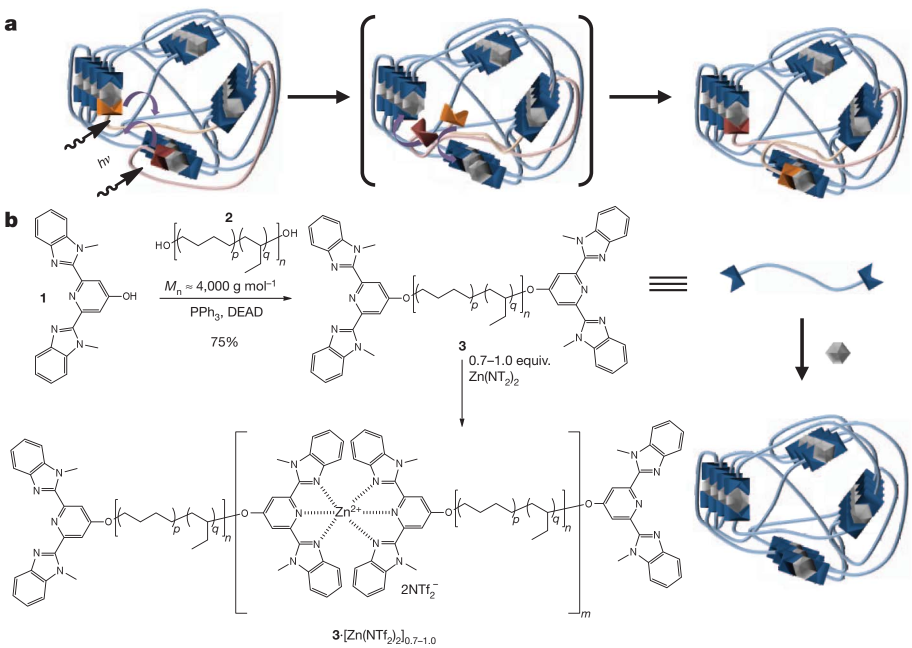


Figure 1 | Mechanism and synthesis of photohealable metallosupramolecular polymers. a, Proposed optical healing of a metallosupramolecular, phase-separated network. **b**, Synthesis of macromonomer **3** and polymerization by addition of $Zn(NTf_2)_2$. DEAD, diethyl azodicarboxylate.

[1]Department of Macromolecular Science and Engineering, Case Western Reserve University, 2100 Adelbert Road, Cleveland, Ohio 44106-7202, USA. [2]US Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005-5069, USA. [3]Adolphe Merkle Institute and Fribourg Center for Nanomaterials, University of Fribourg, CH-1700 Fribourg, Switzerland.

$La(NTf_2)_3$. La^{3+} ions form weaker, more dynamic 3:1 complexes with Mebip than do Zn^{2+} ions, which bind to this ligand in a 2:1 ratio[13,23]. Bistriflimide (NTf_2^{-}) was chosen as the counterion because of its thermal stability and non-coordinating nature. The combination of equimolar amounts of $Zn(NTf_2)_2$ and **3** in solution caused a rapid viscosity increase, indicative of supramolecular assembly (this polymer is denoted as $\mathbf{3}\cdot[Zn(NTf_2)_2]_{1.0}$). Solvent evaporation and compression moulding resulted in colourless elastic films that, unlike **3**, have appreciable mechanical properties. Using the same procedure, polymers with Zn^{2+}:**3** ratios of 0.9–0.7 were also made ($\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.9}$, $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.8}$ and $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$).

Small-angle X-ray scattering (SAXS) and transmission electron microscopy (TEM) studies of films of $\mathbf{3}\cdot[Zn(NTf_2)_2]_x$ revealed microphase-separated lamellar morphologies in which the metal–ligand complexes form a 'hard phase' that physically crosslinks the poly(ethylene-*co*-butylene) 'soft' domains. The SAXS data (Fig. 2a and Supplementary Table 1) show strong Bragg diffraction maxima at integer multiples of the scattering vector magnitude of the primary diffraction peak ($2q^*$, $3q^*$ and so on), characteristic of well-ordered layered morphologies. The lamellar period increased from 8.3 nm for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{1.0}$ to 9.3 nm for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ as the Zn^{2+}:**3** ratio decreased. Concomitantly, the number of strong reflections decreased from four to two, indicating a reduction of long-range order. TEM experiments (Fig. 2b and Supplementary Fig. 2) confirmed the SAXS data.

The thermomechanical properties of **3** and $\mathbf{3}\cdot[Zn(NTf_2)_2]_x$ were probed by modulated differential scanning calorimetry (MDSC; Supplementary Fig. 3) and dynamic mechanical thermal analysis (DMTA; Supplementary Fig. 4). The MDSC trace of **3** shows reversible

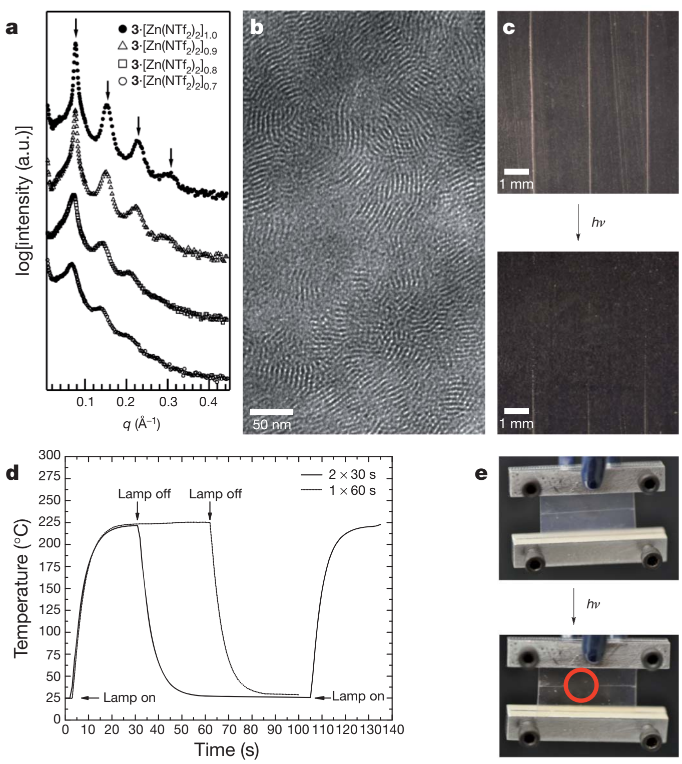


Figure 2 | Characterization of Zn-based metallosupramolecular polymers. **a**, SAXS data for metallosupramolecular polymers of **3** and varying amounts of $Zn(NTf_2)_2$, shifted vertically for clarity. **b**, Representative TEM micrograph showing the lamellar morphology of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{1.0}$. **c**, Optical healing of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ on exposure to light in the wavelength range 320–390 nm for 30 s twice at an intensity of 950 mW cm^{-2}. **d**, Surface temperature of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ on irradiation under the same conditions, as a function of time. **e**, Optical healing of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ while under a load of ~8 kPa (width, 21 mm; thickness, 0.31 mm; mass, 5.25 g).

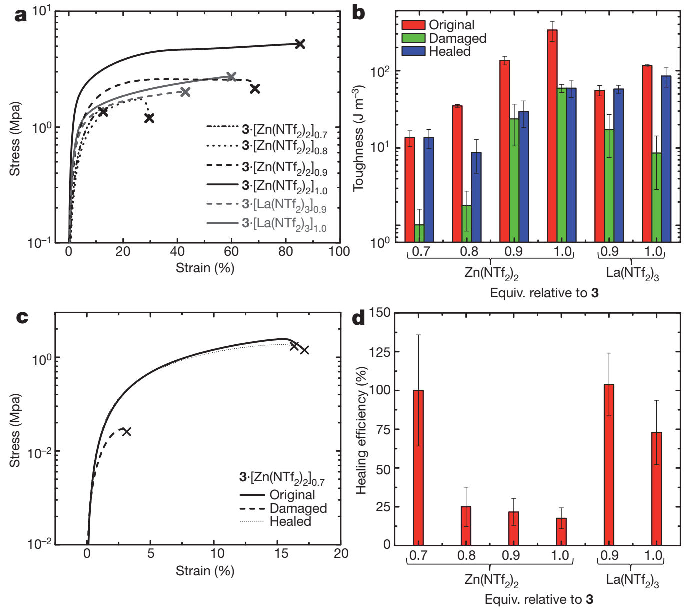


Figure 3 | Mechanical properties and healing of the metallosupramolecular polymers. **a**, Stress–strain curves of metallosupramolecular polymers of **3** and varying amounts of $Zn(NTf_2)_2$ or $La(NTf_2)_3$. **b**, Toughness of metallosupramolecular polymers of **3** and varying amounts of $Zn(NTf_2)_2$ or $La(NTf_2)_3$ (n = 3–5 for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.8}$, $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.9}$, $\mathbf{3}\cdot[Zn(NTf_2)_2]_{1.0}$ and $\mathbf{3}\cdot[La(NTf_2)_3]_{1.0}$; n = 10–13 for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ and $\mathbf{3}\cdot[La(NTf_2)_3]_{0.9}$; error bars, s.d.). **c**, Stress–strain curves of films of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$; shown are data for original, damaged, and healed samples. **d**, Healing efficiency of films of **3** and varying amounts of $Zn(NTf_2)_2$ or $La(NTf_2)_3$. Error bars, s.d.

transitions at −51 and 47 °C, which by comparison with **2** and previous data[16] are assigned as the T_g values of the poly(ethylene-*co*-butylene) core and of the melting of phase-separated Mebip domains, respectively. The metallopolymers $\mathbf{3}\cdot[Zn(NTf_2)_2]_x$ show similar MDSC traces but do not show melting of uncomplexed Mebip domains. However, DMTA traces show the T_g of the poly(ethylene-*co*-butylene) core to be around −23 °C, with a weaker transition at around ~50 °C, consistent with the thermal transition associated with the uncomplexed Mebip domains. At 25 °C, the polymers display storage moduli between 60 and 53 MPa with little dependence on the Zn^{2+}:**3** ratio. The storage modulus drops above ~50 °C for polymers with non-stoichiometric Zn^{2+}:**3** ratios and above 100 °C for all materials, indicative of depolymerization at higher temperatures. Stress–strain experiments conducted at 25 °C show a pronounced decrease in strength, elongation at break and toughness as the Zn^{2+}:**3** ratio is decreased (Fig. 3a, b), consistent with a decrease in molecular mass and long-range order (Fig. 2a, b and Supplementary Fig. 2) as the stoichiometric balance of metal and ligand is offset.

The optical absorption spectrum of **3** has a band with a maximum at 313 nm that is characteristic of Mebip (Supplementary Fig. 5). On complexation of **3** with $Zn(NTf_2)_2$, this band weakens and a new peak, at 341 nm, appears. A titration series shows an isosbestic point at 328 nm, reflecting a well-defined equilibrium between free and metal-coordinated ligands. The absorption spectra of films of **3** and $\mathbf{3}\cdot[Zn(NTf_2)_2]_x$ have similar features (Supplementary Fig. 6). Low-molecular-mass Zn^{2+}–Mebip complexes fluoresce weakly, suggesting that a considerable portion of absorbed light is converted into heat. The optical healing pursued here is based on the assumption that this energy could be harnessed to locally dissociate the supramolecular motif and disengage the macromonomer ends from the hard phase (Fig. 1a), resulting in a decrease in the supramolecular polymer's molecular mass and liquefying the material.

To test this hypothesis, we deliberately damaged 350–400-μm-thick films of the polymers based on **3** and 0.7–1.0 equiv. of $Zn(NTf_2)_2$ by applying well-defined cuts with a depth of ~50–70% of the film thickness. These samples were subsequently exposed to ultraviolet radiation with a wavelength of 320–390 nm and an intensity of 950 mW cm^{-2}. Pictures of films made of $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ suggest that under these conditions two consecutive exposures of 30 s are sufficient to heal the cuts completely (Fig. 2c). Materials with higher metal contents healed less well. Slight discolouration of the samples on ultraviolet irradiation was observed, and increased with exposure time; this may reflect photo-oxidative stress, which is not surprising given the light intensity, the temperature, the presence of air and possibly residual double bonds in the poly(ethylene-*co*-butylene) core. Figure 2e shows that it was readily possible to heal a $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ film while applying a stress of ~8 kPa, by irradiating only the damaged portion. The extinction coefficient of the metallopolymer films is ~890 cm^{-1}, and light absorption and heat generation therefore occur predominantly at the surface. *In situ* experiments revealed that under the conditions employed here for healing, the surface temperature of a $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ film rose to over 220 °C in 30 s and slightly higher on longer irradiation (Fig. 2d). We qualitatively confirmed that $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ films also healed on heating to ~190 °C, which supports the idea that the light-induced healing is indeed due to photothermal conversion. Reference experiments in which $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ films were irradiated with light of similar intensity but a wavelength outside the absorption band of the metal–ligand complex showed no healing, confirming that the process is indeed due to absorption by the metal–ligand complex and not to infrared heating (Supplementary Fig. 7).

To determine the healing efficiency of $\mathbf{3}\cdot[Zn(NTf_2)_2]_x$ in a quantitative manner, films were damaged and healed as described above. We conducted stress–strain experiments on 'original', 'damaged' and 'healed' samples, and the healing efficiency was expressed using toughness as the figure of merit:

$$\text{healing efficiency} = 100\% \frac{\text{toughness}_{\text{healed}}}{\text{toughness}_{\text{original}}}$$

Here $\text{toughness}_{\text{healed}}$ is the toughness of the healed sample set and $\text{toughness}_{\text{original}}$ is the toughness of the original sample set.

Figure 3b shows that the strain and stress at break of a $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$ film (and, hence, toughness, which is proportional to the area under the stress–strain curve) are significantly reduced on damaging the samples as described above. However, the original properties could be restored on ultraviolet irradiation. Statistical experiments (Fig. 3b, d and Supplementary Tables 2 and 3) show that $\text{toughness}_{\text{healed}}$ is comparable to $\text{toughness}_{\text{original}}$ and statistically different from the toughness of the damaged sample set, with a healing efficiency of $100 \pm 36\%$. Lower healing efficiencies, of $25 \pm 12\%$, $22 \pm 8.6\%$ and $18 \pm 6.7\%$, were observed for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.8}$, $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.9}$ and $\mathbf{3}\cdot[Zn(NTf_2)_2]_{1.0}$, respectively. This is consistent with the greatest viscosity reduction being for $\mathbf{3}\cdot[Zn(NTf_2)_2]_{0.7}$, on account of the excess of free ligands, which render the system more dynamic and lower the viscosity. However, the concomitant decrease in molecular mass and long-range order also lower the toughness as the Zn^{2+}:**3** ratio is decreased (Fig. 3a, b).

As La^{3+}–Mebip complexes are more labile and dissociate at lower temperatures than Zn^{2+}–Mebip complexes[13], we also studied polymers based on **3** and 1.0 or 0.9 equiv. of $La(NTf_2)_3$ ($\mathbf{3}\cdot[La(NTf_2)_3]_{1.0}$ and $\mathbf{3}\cdot[La(NTf_2)_3]_{0.9}$). Titration experiments confirmed that the La–$MeBip(NTf_2)_3$ complexes involve 1:3 binding (Supplementary Fig. 9), suggesting the formation of network structures. Nevertheless, SAXS and TEM data reveal the formation of a predominantly lamellar morphology (Fig. 4a, b) with a period of 7.4 nm (SAXS data for

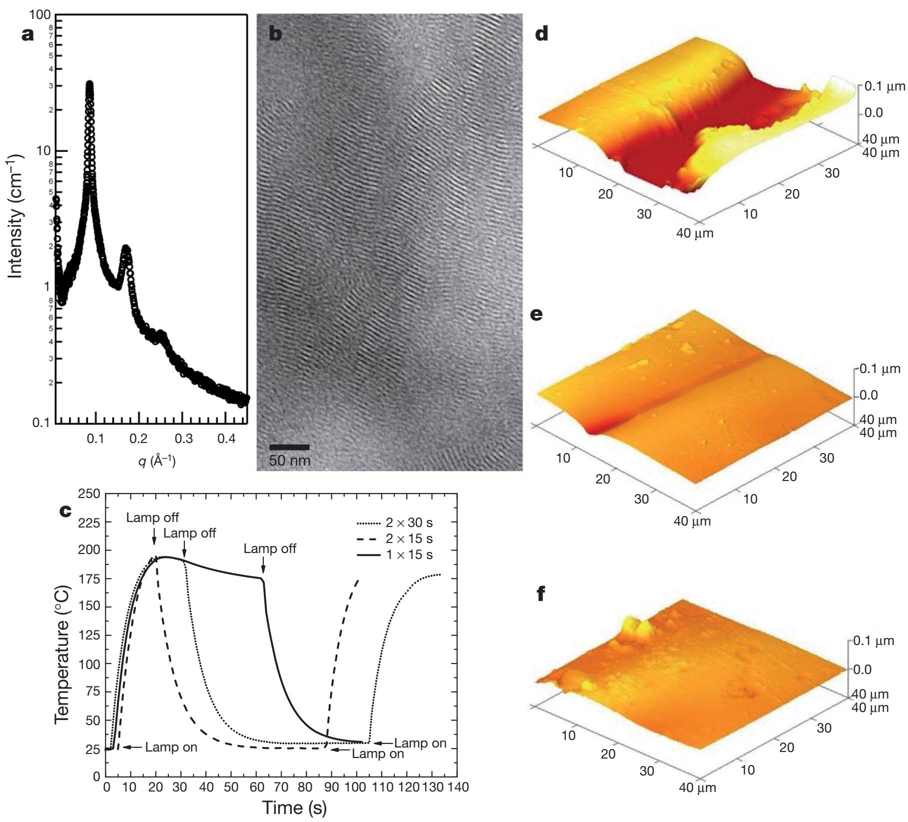


Figure 4 | Characterization of La-based metallosupramolecular polymers. **a**, SAXS data for metallosupramolecular polymers of $\mathbf{3}\cdot[La(NTf_2)_3]_{1.0}$. **b**, Representative TEM micrograph showing the lamellar morphology observed in $\mathbf{3}\cdot[La(NTf_2)_3]_{1.0}$. **c**, Temperature of a $\mathbf{3}\cdot[La(NTf_2)_3]_{0.9}$ film on exposure to healing conditions, measured using an infrared camera. **d–f**, AFM images of a damaged $\mathbf{3}\cdot[La(NTf_2)_3]_{1.0}$ film before healing (**d**), after partial healing (**e**) and after complete healing (**f**).

$3\cdot[La(NTf_2)_3]_{1.0}$). The mechanical properties of $3\cdot[La(NTf_2)_3]_x$ are similar to those of $3\cdot[Zn(NTf_2)_2]_{0.9}$ (Fig. 3 and Supplementary Fig. 11), but their healing efficiencies, of $104 \pm 20\%$ and $73 \pm 21\%$, for $3\cdot[La(NTf_2)_3]_{0.9}$ and $3\cdot[La(NTf_2)_3]_{1.0}$, respectively (Fig. 3b, d), are much better. The surface temperature of $3\cdot[La(NTf_2)_3]_{0.9}$ during photohealing, of $\sim 195\,^{\circ}C$ (Fig. 4c), was lower than that of $3\cdot[Zn(NTf_2)_2]_{0.7}$, and optical microscopy showed qualitatively that the former liquefied more readily, consistent with the weaker and more dynamic binding. Thus, $3\cdot[La(NTf_2)_3]_{0.9}$ shows the best combination of high toughness and high healing efficiency, which seems to be related to the more dynamic nature of La–MeBip complexes relative to Zn–MeBip complexes. We used atomic force microscopy (AFM) to probe the topology of a $3\cdot[La(NTf_2)_3]_{1.0}$ film around a defect before and after healing (Fig. 4d–f). The images show that the cut is filled and disappears, consistent with the proposed viscosity decrease on exposure to light.

Our systematic investigation of several strategically chosen compositions and in-depth morphological studies provide insight into the healing process in these metallosupramolecular polymers, based on a previously unexplored combination of supramolecular polymerization and the concept of light–heat conversion. The formation of lamellar morphologies in which a hard phase comprising the metal–ligand complexes physically crosslinks soft domains of the poly(ethylene-*co*-butylene) cores is the main determinant for the thermomechanical characteristics of the materials studied. The data suggest that the dynamics of the light-induced depolymerization and, thereby, the healing behaviour are governed by the presence of an excess of free ligands and the nature of the metal–ligand bond. The concept of photothermally induced healing of supramolecular materials seems to be applicable to any supramolecular polymer with a binding motif that is sufficiently dynamic. The ability to change the chromophore makes it possible to tailor the wavelength required for healing. The combination of the new approach with an additional mechanochromic response[24,25] promises access to true—that is, autonomously functioning—self-healing materials, in which light is absorbed only at defect sites.

METHODS SUMMARY

We synthesized macromonomer **3** (number-average molecular mass, $M_n = 4{,}800\,g\,mol^{-1}$) via a Mitsunobu reaction of **1** and hydroxyl-terminated poly (ethylene-*co*-butylene) (**2**, $M_n = 4{,}000\,g\,mol^{-1}$). $3\cdot[Zn(NTf_2)_2]_x$ and $3\cdot[La(NTf_2)_3]_x$ were formed by combining **3** with $Zn(NTf_2)_2$ or $La(NTf_2)_3$ in a CH_3Cl/CH_3CN mixture. Films 350–400 µm thick were produced by casting and drying these solutions and compression moulding the resulting solids at $120\,^{\circ}C$. The thermomechanical properties were characterized by MDSC and DMTA under N_2 at heating rates of $3\,^{\circ}C\,min^{-1}$. Stress–strain experiments were conducted at $25\,^{\circ}C$ with a strain rate of $5\%\,min^{-1}$, on 'dog-bone-shaped' samples (Supplementary Fig. 8). Damage to the film was applied by cutting the samples to a depth of ~50–70% with a razor blade. Samples were healed by exposure to light of wavelength 320–390 nm and intensity $950\,mW\,cm^{-2}$. Mechanical data are averages of $n = 3$–5 independent experiments for $3\cdot[Zn(NTf_2)_2]_{0.8}$, $3\cdot[Zn(NTf_2)_2]_{0.9}$, $3\cdot[Zn(NTf_2)_2]_{1.0}$ and $3\cdot[La(NTf_2)_3]_{1.0}$, and $n = 10$–13 independent experiments for $3\cdot[Zn(NTf_2)_2]_{0.7}$ and $3\cdot[La(NTf_2)_3]_{0.9}$. All errors are standard deviations. On account of the destructive nature of the stress–strain experiments, different samples from the same batch were used to measure the materials in their original, damaged and healed states. To account for the thickness variation between samples, force–displacement curves were divided by the nominal cross-sectional area of the sample (assuming no defect for damaged and healed samples) to yield stress–strain curves. These were integrated to yield toughness as a figure of merit, allowing comparison of original, damaged and healed samples. We characterized morphology using SAXS, TEM and AFM. Surface temperatures were determined with an infrared camera.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 September 2010; accepted 15 February 2011.

1. Blaiszik, B. J. *et al.* Self-healing polymers and composites. *Annu. Rev. Mater. Res.* **40,** 179–211 (2010).
2. Bergman, S. D. & Wudl, F. Mendable polymers. *J. Mater. Chem.* **18,** 41–62 (2008).
3. Wool, R. P. Self-healing materials: a review. *Soft Matter* **4,** 400–418 (2008).
4. Murphy, E. B. & Wudl, F. The world of smart healable materials. *Prog. Polym. Sci.* **35,** 223–251 (2010).
5. Cordier, P., Tournilhac, F., Soulie-Ziakovic, C. & Leibler, L. Self-healing and thermoreversible rubber from supramolecular assembly. *Nature* **451,** 977–980 (2008).
6. Kim, Y. H. & Wool, R. P. A theory of healing at a polymer-polymer interface. *Macromolecules* **16,** 1115–1120 (1983).
7. Chen, X. *et al.* A thermally re-mendable cross-linked polymeric material. *Science* **295,** 1698–1702 (2002).
8. Murphy, E. B. *et al.* Synthesis and characterization of a single-component thermally remendable polymer network: Staudinger and Stille revisted. *Macromolecules* **41,** 5203–5209 (2008).
9. Burattini, S. *et al.* A self-repairing, supramolecular polymer system: healability as a consequence of donor-acceptor pi-pi stacking interactions. *Chem. Commun. (Camb.)*6717–6719 (2009).
10. Burattini, S. *et al.* A healable supramolecular polymer blend based on aromatic $\pi-\pi$ stacking and hydrogen-bonding interactions. *J. Am. Chem. Soc.* **132,** 12051–12058 (2010).
11. Wojtecki, R. J., Meador, M. A. & Rowan, S. J. Utilizing the dynamic bond to access macroscopically-responsive structurally-dynamic polymers. *Nature Mater.* **10,** 14–27 (2011).
12. Bosman, A. W., Sijbesma, R. P. & Meijer, E. W. Supramolecular polymers at work. *Mater. Today* **7,** 34–39 (2004).
13. Kumpfer, J. R., Jin, J. Z. & Rowan, S. J. Stimuli-responsive europium-containing metallo-supramolecular polymers. *J. Mater. Chem.* **20,** 145–151 (2010).
14. Sivakova, S., Bohnsack, D. A., Mackay, M. E., Suwanmala, P. & Rowan, S. J. Utilization of a combination of weak hydrogen-bonding interactions and phase segregation to yield highly thermosensitive supramolecular polymers. *J. Am. Chem. Soc.* **127,** 18202–18211 (2005).
15. Ghosh, B. & Urban, M. W. Self-repairing oxetane-substituted chitosan polyurethane networks. *Science* **323,** 1458–1460 (2009).
16. Kautz, H., van Beek, D. J. M., Sijbesma, R. P. & Meijer, E. W. Cooperative end-to-end and lateral hydrogen-bonding motifs in supramolecular thermoplastic elastomers. *Macromolecules* **39,** 4265–4267 (2009).
17. Beck, J. B., Ineman, J. M. & Rowan, S. J. Metal/ligand-induced formation of metallo-supramolecular polymers. *Macromolecules* **38,** 5060–5068 (2005).
18. Burnworth, M., Knapton, D., Rowan, S. J. & Weder, C. Metallo-supramolecular polymerization: a route to easy-to-process organic/inorganic hybrid materials. *J. Inorg. Organomet. Polym. Mater.* **17,** 91–103 (2007).
19. Burnworth, M., Mendez, J. D., Schroeter, M., Rowan, S. J. & Weder, C. Decoupling optical properties in metallo-supramolecular poly(*p*-phenylene ethynylene)s. *Macromolecules* **41,** 2157–2163 (2008).
20. Knapton, D., Rowan, S. J. & Weder, C. Synthesis and properties of metallo-supramolecular poly(*p*-phenylene ethynylene)s. *Macromolecules* **39,** 651–657 (2006).
21. Knapton, D., Iyer, P. K., Rowan, S. J. & Weder, C. Synthesis and properties of metallo-supramolecular poly(*p*-xylylene)s. *Macromolecules* **39,** 4069–4074 (2006).
22. Knapton, D., Burnworth, M., Rowan, S. J. & Weder, C. Fluorescent organometallic sensors for the detection of chemical-warfare-agent mimics. *Angew. Chem. Int. Ed.* **45,** 5825–5829 (2006).
23. Beck, J. B. & Rowan, S. J. Metal-ligand induced supramolecular polymerization: a route to responsive materials. *Faraday Discuss.* **128,** 43–53 (2005).
24. Kunzelman, J., Kinami, M., Crenshaw, B. R., Protasiewicz, J. D. & Weder, C. Oligo(*p*-phenylene vinylene)s as a ''new'' class of piezochromic fluorophores. *Adv. Mater.* **20,** 119–122 (2008).
25. Crenshaw, B. R. *et al.* Deformation-induced color changes in mechanochromic polyethylene blends. *Macromolecules* **40,** 2400–2408 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This material is based on work supported by the US Army Research Office (W911NF-09-1-0288 and W911NF-06-1-0414); the National Science Foundation under grant numbers CHE-0704026, DMR-0602869 and MRI-0821515; the Adolphe Merkle Foundation; and the Postgraduate Research Participation Program at the US Army Research Laboratory, administered by the Oak Ridge Institute of Science and Education through an interagency agreement between the US Department of Energy and Army Research Laboratory (contract number ORISE-1120-1120-99). We thank S. Dellinger for the design and the fabrication of a device to introduce well-defined defects and Kraton Performance Polymers Inc for the donation of the hydroxyl-terminated poly(ethylene-*co*-butylene).

Author Contributions M.B., L.T. and J.R.K. developed the procedures for synthesis and characterization of **3**. M.B. prepared and processed all supramolecular polymers. M.B. and G.L.F. did the MDSC experiments. A.J.D. and F.L.B. carried out the TEM and SAXS experiments. M.B. did the mechanical testing. G.L.F. carried out the light–heat conversion experiments. M.B. conducted the photohealing experiments. S.J.R. and C.W. designed the study. All authors discussed results and contributed to the interpretation of data. M.B., S.J.R. and C.W. wrote the paper. All authors contributed to editing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.J.R. (stuart.rowan@case.edu) or C.W. (christoph.weder@unifr.ch).

METHODS

General methods. ^1H and ^{13}C NMR spectra were acquired in $CDCl_3$ using a Varian 600-MHz spectrometer; chemical shifts (δ) are expressed in parts per million relative to an internal TMS standard. Mass spectra were obtained using a Bruker AUTOFLEX III MALDI TOF/TOF mass spectrometer using dithranol as a matrix. Size exclusion chromatography (SEC) of macromonomer **3** was conducted on a Visotek Model 270 Dual Detector, equipped with Varian OligoPore and ResiPore SEC/GPC columns. Ultraviolet–visible absorption spectra were obtained on a Perkin-Elmer Lambda 800 spectrometer. MDSC experiments were carried out under nitrogen on TA Instrument DSCQ2000 at a heating rate of 3 °C min^{-1}. DMTA measurements were conducted under nitrogen on a Perkin-Elmer Q800 spectrometer at a heating rate of 3 °C min^{-1}. Stress–strain experiments were conducted on the same instrument at a temperature of 25 °C and a strain rate of 5% min^{-1}. Surface temperatures of thin films were determined with an Optris PI Connect infrared camera from Roth and Co AG (Oberuzwil).

Synthesis of macromonomer 3. To a stirred solution of 2,6-bis(1′-methylbenzimidazolyl)-4-hydroxypyridine[26] (2.30 g, 6.47 mmol), triphenylphosphine (3.40 g, 13.0 mmol), hydroxyl-terminated poly(ethylene-*co*-butylene) (M_n = 4,000 g mol^{-1}, 9.1 g, 2.0 mmol; donated by Kraton) in dry THF (120 ml), 5 ml of diethyl azodicarboxylate (DEAD) 40 wt% in toluene was added at −40 °C. After 4 h the reaction mixture was allowed to warm to 25 °C and was left to react at this temperature for 44 h. The solvents were then evaporated in vacuum and the remaining residue was redissolved in hot hexanes (150 ml) and washed three times with MeOH and three times with 1 M NaOH(aq.). The product was purified by column chromatography ($CHCl_3$, SiOH), dried overnight and collected as a tacky white solid (9.0 g, 81% yield). ^1H NMR (600 MHz, $CDCl_3$, 18 °C): δ = 7.930 (s, 4H), 7.871 (d, J = 7.8 Hz, 4H), 7.463 (d, J = 7.8 Hz, 4H), 7.382 (m, 4H), 7.351 (m, 4H), 4.239 (s, 12H), 4.252 (t, 4H), 1.480–0.935 (m, 438H), 0.935–0.79 (m, 138H). ^{13}C NMR (100 MHz, $CDCl_3$, 18 °C): δ = 166.67, 151.01, 150.45, 142.47, 137.15, 123.50, 122.78, 120.13, 111.78, 109.89, 68.680, 39.040–37.808, 36.039, 33.380, 33.190, 32.531, 30.650, 30.194, 29.738, 26.745–25.822, 10.860–10.196. As determined by NMR, X_n = 72 (number-average degree of polymerization) and M_n = 4,800 g mol^{-1}. As determined by SEC, M_n = 6,950 g mol^{-1}, M_w = 7,080 g mol^{-1} (weight average) and PDI = 1.02 (polydispersity index).

Supramolecular polymerization and film formation. To a stirred solution of **3** in $CHCl_3$, ~0.025 M, a solution of $Zn(NTf_2)_2$ or $La(NTf_2)_3$ in CH_3CN, ~0.05 M, was added. The amount of metal added was adjusted to provide for metal:**3** equivalents of 1.0–0.7 (accounting for 1:2 and 1:3 binding, respectively), within an error of ~5%. On addition of the metal salt, a white precipitate formed. This precipitate subsequently dissolved in 15 s for the Zn^{2+} systems and an increase of the solutions viscosity was observed. The La^{3+} systems led to homogenous gels after 5 min of stirring. The solutions were then dried in vacuum to produce the supramolecular polymers **3**·$[Zn(NTf_2)_2]_x$ and **3**·$[La(NTf_2)_3]_x$, which were subsequently compression moulded at 120 °C and a pressure of 3 t for 10 min to produce 350–400-μm-thick films. To conduct optical absorption experiments, films of a thickness of 2.7–6.3 μm were prepared by dissolving the polymers in $CHCl_3$ (concentration, 25 mg ml^{-1}) and spin coating.

Mechanical testing. The mechanical properties of films of **3**·$[Zn(NTf_2)_2]_x$ and **3**·$[La(NTf_2)_3]_x$ were tested by DMTA using either a controlled-frequency temperature sweep or a controlled strain rate at constant temperature. The temperature sweeps were conducted under N_2 with a frequency of 1 Hz and a heating rate of 3 °C min^{-1} from −100 °C to 100 °C. Stress–strain data were collected under N_2 at 25 °C, with a strain rate of 5% min^{-1}. All mechanical tests were conducted on dog-bone-shaped samples (Supplementary Fig. 8). Different samples from the same batch were used to measure the materials in their original, damaged and healed states.

Film damaging and healing. Films of **3**·$[Zn(NTf_2)_2]_x$ and **3**·$[La(NTf_2)_3]_x$ were damaged by cutting the samples to a depth of ~50–70% of their thickness with a razor blade attached to a calliper, which allowed for specific depth control. These samples were then exposed to a filtered light source (Bluepoint 4 Ecocure from Honle UV America), which irradiated the samples with wavelengths of light from 320 to 390 nm at an intensity of 950 mW cm^{-2}. Reference experiments were conducted with the same lamp system, but with a filter that changed the wavelength to 400–500 nm.

Morphology characterization. SAXS data were collected using a Molecular Metrology 120-mm-diameter detector, Ni-filtered $Cu_{K\alpha}$ X-ray radiation (λ = 1.542 Å) produced using a Rigaku Ultrax18 X-ray generator, and sample-to-detector distances of 150 and 50 cm. The instrument was calibrated using Ag behenate and glassy carbon standards. The data were corrected for sample absorption and background noise before scaling to absolute intensity (cm^{-1}). The two-dimensional data sets were azimuthally averaged to give intensity as a function of the magnitude of the scattering vector, q, where $q = 4\pi\sin(\theta)/\lambda$ and 2θ is the scattering angle. All data processing and analysis were performed using IGOR PRO 6 (WaveMetrics) and tool suites developed in ref. 27. Bragg diffraction maxima were fitted using Lorentzian distributions. TEM was performed using a JEOL JEM-2100F field-emission TEM operated at 200 kV. Sections approximately 50–70-nm thick were prepared for TEM by ultramicrotomy using a Leica UCT ultracryomicrotome. Bright-field images were collected with a Gatan SC1000 Orius 11-megapixel charge-coupled-device camera and analysed using Gatan DIGITALMICROGRAPH software. AFM data were collected on a Veeco MultiMode V AFM run in tapping mode.

26. Beck, J. B. & Rowan, S. J. Multistimuli, multiresponsive metallo-supramolecular polymers. *J. Am. Chem. Soc.* **125,** 13922–13923 (2003).
27. Ilavsky, J. & Jemian, P. R. *Irena*: tool suite for modelling and analysis of small angle scattering. *J. Appl. Crystallogr.* **42,** 347–353 (2009).

doi:10.1038/nature09903

Thermal history of Mars inferred from orbital geochemistry of volcanic provinces

David Baratoux[1,2], Michael J. Toplis[1,2], Marc Monnereau[1,2] & Olivier Gasnault[1,2]

Reconstruction of the geological history of Mars has been the focus of considerable attention over the past four decades, with important discoveries being made about variations in surface conditions[1]. However, despite a significant increase in the amount of data related to the morphology, mineralogy and chemistry of the martian surface, there is no clear global picture of how magmatism has evolved over time and how these changes relate to the internal workings and thermal evolution of the planet. Here we present geochemical data derived from the Gamma Ray Spectrometer on board NASA's Mars Odyssey spacecraft[2], focusing on twelve major volcanic provinces of variable age. Our analysis reveals clear trends in composition that are found to be consistent with varying degrees of melting of the martian mantle. There is evidence for thickening of the lithosphere (17–25 km Gyr^{-1}) associated with a decrease in mantle potential temperature over time (30–40 K Gyr^{-1}). Our inferred thermal history of Mars, unlike that of the Earth, is consistent with simple models of mantle convection[3–6].

Partial melting of planetary interiors leads to the formation of liquids whose composition is directly related to the pressure–temperature conditions of melting. If such liquids reach the surface, for example as lava flows, this opens the possibility of using surface composition to infer degrees and depths of melting as functions of time. The link between magma chemistry and mantle thermal state has been used to rationalize certain variations in terrestrial magmatism[7], but has never been applied to other planetary bodies, principally because of a lack of appropriate data.

The Gamma Ray Spectrometer (GRS) instrument provides the first opportunity to explore this possibility elsewhere than on the Earth, as the abundance of the rock-forming elements silicon, iron, potassium, thorium and chlorine has been quantified at the martian surface over an equatorial band extending from ~45° north to ~45° south[2]. Consideration of the entire GRS data set has been used to identify regions of similar chemistry[8,9], although detailed interpretation of these provinces is complicated by the fact that surface chemistry may be affected by primary magmatism, subsequent chemical alteration and transport processes. To highlight magmatic signatures, an alternative approach is to restrict analysis to regions with clear morphological evidence for a volcanic origin[10] (for example lava flows, volcanic vents, calderas or shield volcanoes). Regions have been selected that are significantly larger than the GRS footprint and which fall well within the larger units previously defined on the basis of GRS data[9]. As such, average chemical composition is independent of the exact shape and size of each region. With these criteria, we retained 12 provinces (Fig. 1), six of Hesperian age (3.6–4.0 Gyr) and six of more recent, Amazonian, age.

Our analysis focuses on iron, silicon and thorium. We did not explicitly consider potassium as variations of K/Th are remarkably modest[11] and thorium is known to be less mobile than potassium during aqueous alteration. Chlorine was not considered as although it may be of magmatic origin[12], it is also known to be concentrated in secondary salts[13]. Average and standard deviations of FeO (parts per hundred by weight (wt%)), SiO_2 (wt%) and thorium (parts per million by weight (wt p.p.m.)) in each region were calculated from GRS data released to NASA's Planetary Data System[2] and renormalized on a hydrogen-free basis (Table 1).

All three elements show significant variability and a prominent anticorrelation is observed between SiO_2 and thorium (Fig. 2a), with younger volcanoes being characterized by lower SiO_2 and higher thorium contents (Fig. 2a). Values of FeO are generally lower for the younger edifices of the Tharsis dome (Fig. 2b). Comparison of these trends with possible global dust compositions argues against the possibility that geochemical variations result from a variable fraction of transported material (Supplementary Information).

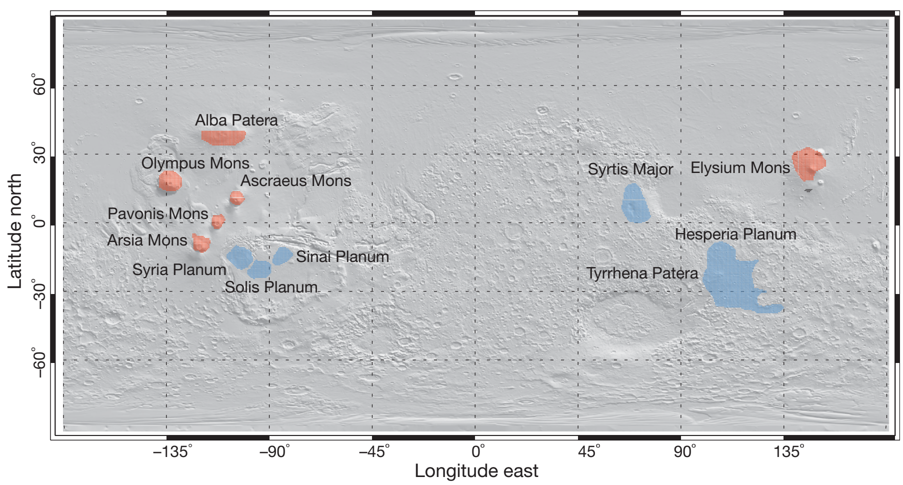


Figure 1 | Selected volcanic provinces represented over a shaded relief map of Mars. Hesperian volcanic provinces, located in the highlands south of the hemispheric dichotomy, are highlighted in blue and Amazonian provinces occurring as large shield volcanoes are highlighted in red.

[1]Université de Toulouse, UPS-OMP, IRAP, F-31400 Toulouse, France. [2]CNRS, IRAP, 14 Avenue Edouard Belin, F-31400 Toulouse, France.

Table 1 | Average composition and s.d. for each volcanic province

Volcanic province	SiO_2 (wt%)	FeO (wt%)	Th (wt p.p.m.)	Age (period)
Olympus Mons	42.0 ± 0.6	17.1 ± 0.7	0.58 ± 0.02	Amazonian
Arsia Mons	41.9 ± 0.1	15.7 ± 0.5	0.55 ± 0.03	Amazonian
Ascraeus Mons	43.4 ± 0.2	19.5 ± 0.2	0.62 ± 0.02	Amazonian
Pavonis Mons	42.1 ± 0.2	18.7 ± 0.4	0.60 ± 0.02	Amazonian
Alba Patera	42.3 ± 0.3	19.2 ± 0.6	0.51 ± 0.03	Amazonian
Elysium Mons	43.5 ± 0.4	21.5 ± 0.3	0.49 ± 0.03	Amazonian
Syrtis Major	44.7 ± 0.3	20.5 ± 0.9	0.58 ± 0.02	Hesperian
Tyrrhena Patera	45.2 ± 0.2	18.6 ± 0.1	0.43 ± 0.03	Hesperian
Hesperia Planum	45.5 ± 0.5	18.9 ± 0.4	0.48 ± 0.02	Hesperian
Syria Planum	44.4 ± 0.02	18.3 ± 0.5	0.58 ± 0.02	Hesperian
Solis Planum	46.4 ± 0.6	18.5 ± 0.4	0.38 ± 0.04	Hesperian
Sinai Planum	47.3 ± 0.3	19.3 ± 0.2	0.36 ± 0.03	Hesperian

All values are normalized to water-free composition.

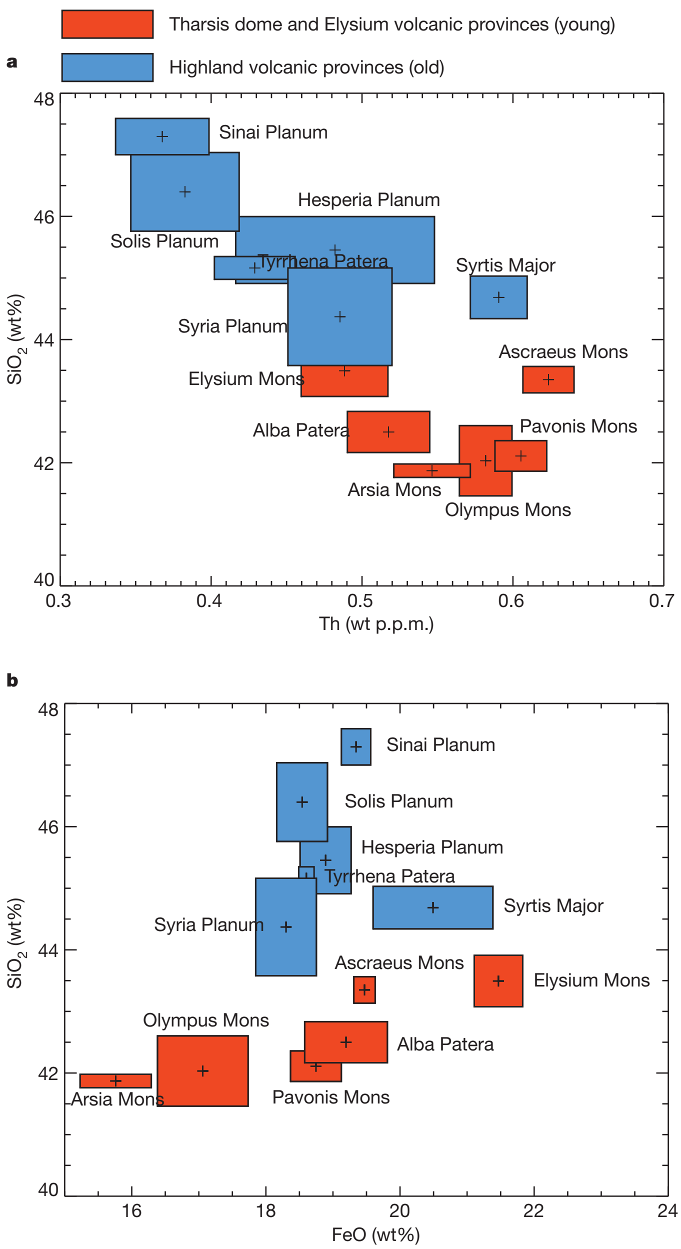


Figure 2 | Abundances of silica, thorium and iron oxide for the 12 selected areas shown in Fig. 1. a, Th versus SiO_2. **b**, FeO versus SiO_2. Older Hesperian volcanoes are shown in blue and younger volcanic provinces are shown in red. Box sizes correspond to errors in element concentration of 1σ.

Assuming a magmatic origin, interpretation of these variations requires an estimate of the mantle composition and an understanding of the magmatic processes leading to volcanic eruption. Concerning mantle chemistry, constraints from chondritic and shergottite–nakhlite–chassignite[14] meteorites have been used to estimate the composition of the primitive martian mantle[15], a composition henceforth referred to as DW85. Although there is some debate surrounding the question of whether or not this composition is the source of shergottites[16,17] the DW85 composition has been shown to be consistent with Tharsis volcanism[10] and, most importantly, with basaltic rocks analysed *in situ* at Gusev crater[18].

In terms of magmatic processes, we assume that melts are produced by adiabatic decompression of the DW85 composition associated with convective upwelling and that liquids equilibrate with the mantle at a single pressure and temperature at the base of the lithosphere. We begin by considering the simplest case, of no fractional crystallization on the way to the surface. Although primary magmas are rare on the Earth, consideration of magnesium number (Mg#, defined as molar Mg/(Mg + Fe) × 100) suggests that they may be much more common at the surface of Mars[19]. For example, whole-rock data from Gusev crater show that the spatially dominant basalts have Mg# ranging from 51 to 55 (ref. 19), consistent with that calculated for primary melts in equilibrium with the DW85 mantle (49 ± 4). Furthermore, the olivine in Shergottite meteorites has core compositions with Mg# typically of 70 to 80, consistent with the range expected for olivine crystallizing from primary magmas derived from a DW85 source. Although these comparisons support the idea of limited degrees of fractional crystallization, we also note that our derived pressures of melting are insensitive to <15% of olivine removal (Methods and Supplementary Information).

We determined optimal values of pressure (P) and the degree of partial melting (F) for each volcanic province by comparison with primary melt compositions calculated using the pMELTS software[20] (Methods). Pressure is generally well constrained, but a significant range of F may explain the observed silica/iron oxide abundances (Methods and Supplementary Information). However, an independent assessment of F is possible using thorium concentrations. Being highly incompatible, the concentration of this element in the liquid is inversely proportional to the degree of partial melting. Assuming the thorium value of DW85 (0.056 wt p.p.m.), values of F derived from silica–iron systematics and from thorium concentrations are in excellent agreement for the six Hesperian provinces. This leads to the conclusion that the DW85 mantle composition is relevant to Hesperian volcanism and that a model of direct transfer of magma from the mantle to the surface is sufficient. However, for younger volcanism values of SiO_2, FeO and thorium do not always point to a consistent P–F solution.

The effects of low-pressure fractional crystallization, assimilation of cumulates, variations in the Mg# of the mantle and thorium depletion have all been tested (Methods and Supplementary Information). Of these, only the last two can potentially account for the Amazonian data. The concentration of highly incompatible elements in the mantle is sensitive to even small amounts of previous melting (Methods and Supplementary Information). For this reason, the hypothesis of a modest depletion of thorium in the Amazonian mantle relative to the value of DW85 has been retained.

The results (Fig. 3) show that older volcanism is characterized by lower P and higher F values relative to younger volcanism (typically >10% melting at pressures of <1.5 GPa for older volcanism and <10% melting at pressures of >1.5 GPa for younger volcanism). As an independent check, we used *in situ* measurements of iron and silicon concentrations of the Hesperian-age Adirondack-class basalts in Gusev crater[18,19] to calculate P and F (Fig. 3). These derived values are consistent with those determined from experimental phase equilibria (15–20% melting at 1.4 GPa (ref. 18)) and are comparable to those of our Hesperian volcanic group.

The variations in P and F imply a change in the thermal state of the mantle that may be quantified by consideration of potential

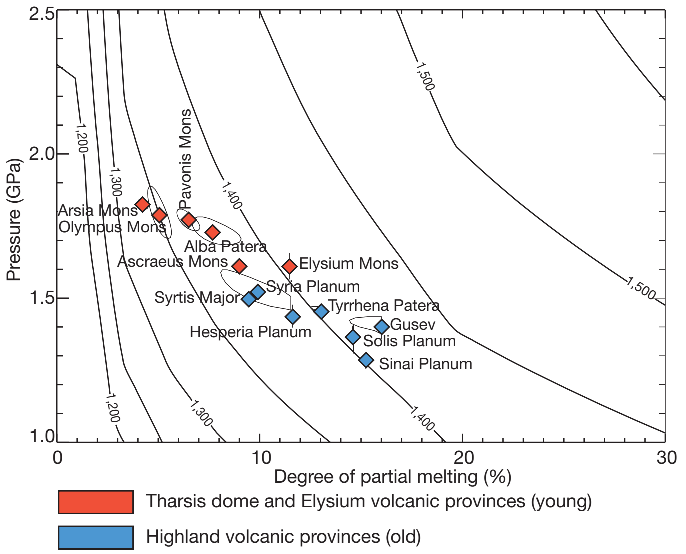


Figure 3 | Pressure and degree of partial melting for each volcanic province. Pressures are derived from SiO_2 and FeO abundances (Methods). In general, F is determined from thorium content in the DW85 mantle except when inconsistent with values derived from SiO_2 and FeO. In these cases, F values derived from SiO_2 and FeO concentrations are given, retaining the hypothesis of a mantle depleted in thorium (Supplementary Information). 1σ error bars and ellipses correspond to propagated uncertainties in SiO_2 and FeO. Contours correspond to mantle potential temperature in degrees Celsius, calculated by adding the latent heat of fusion (equivalent to ~4 K for each per cent of melting) to, and subtracting the effect of adiabatic decompression to surface pressure from, the equilibrium temperature at which liquid and solid mantle coexist[21].

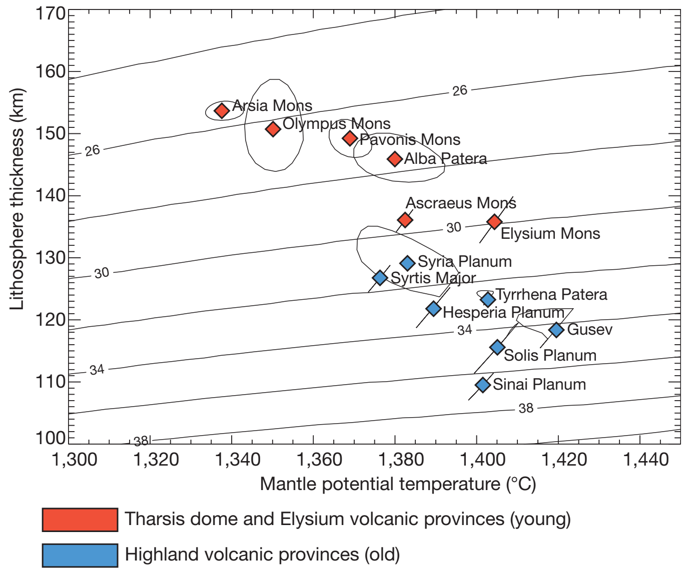


Figure 4 | Lithosphere thickness as a function of mantle potential temperature estimated from pressure and degree of partial melting. Heat flow contours (milliwatts per square metre) are calculated from the lithosphere thickness and the melting temperature assuming a conductivity of 3.5 W m^{-1} K^{-1}. 1σ error bars and ellipses are as calculated in Fig. 3.

temperature[21] (that is, the temperature that unmolten mantle would have if it were taken adiabatically to a pressure of 1 bar). Our data indicate that older volcanoes are produced from a hotter mantle than younger ones, with a total spread of ~80 K (Fig. 3). This cooling is associated with thickening of the lithosphere, from 100 to ~150 km (Fig. 4), implying a heat flow decrease over time from 38 to 26 mW m^{-2} (Fig. 4). These values are in agreement with those derived independently from the thickness of the elastic lithosphere required to support major topographic features: 20–25 mW m^{-2} for the young Tharsis Montes, ~30 mW m^{-2} for Alba Patera and Elysium Mons, and 30–40 mW m^{-2} for the Hesperian volcanism[22,23]. Furthermore,

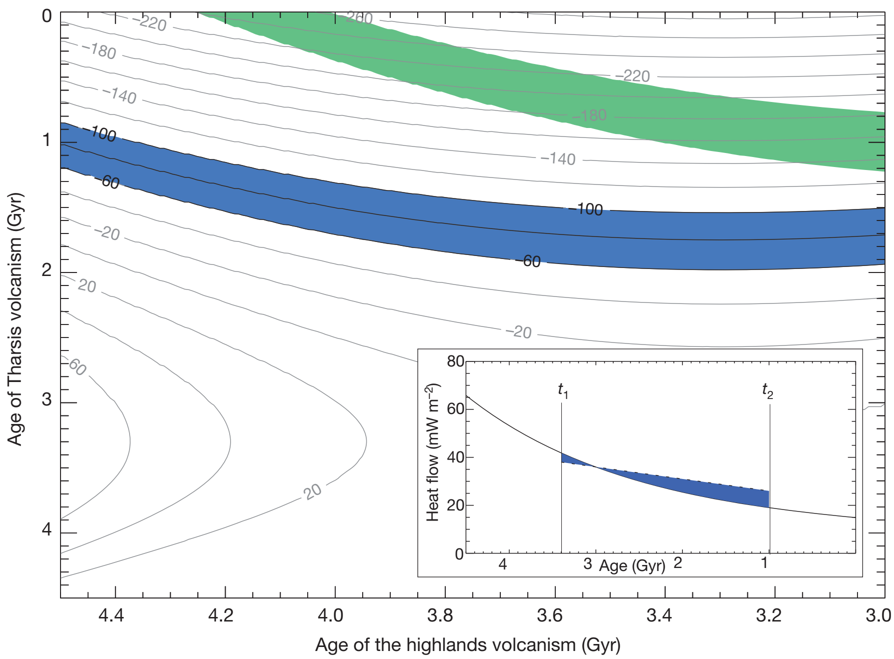


Figure 5 | Mantle temperature variation as a function of the ages of Hesperian volcanism (t_1) and Amazonian volcanism (t_2). The inset illustrates this calculation for the variation of heat flow as a function of time. In more detail, temperatures are calculated from integration between t_1 and t_2 of the difference between heat production in the mantle (solid line) and the inferred heat flow decrease from 36 to 28 mW m^{-2} (dotted line), divided by the heat capacity of the planet. Two possibilities for the distribution of heat-producing elements in the mantle are considered. For the first, values for heat-producing elements are those of DW85. For the second, a depleted mantle composition has been calculated assuming a value of 5.17×10^{23} kg for the mantle mass[29], a crustal mass of 1.74×10^{22} kg (corresponding to an average crustal thickness of 46 km derived from MOLA data[30]), average potassium and thorium abundances from GRS data[2], and a Th/U ratio of 3.6 (ref. 15). Calculated temperature variations (degrees Celsius) are shown as grey contours for the case of depleted mantle. Acceptable solutions, corresponding to a temperature decrease of 80 ± 20 °C, have been highlighted in blue for the depleted mantle and in green for the primitive mantle.

the geochemically derived decrease of 12 mW m^{-2} associated with a drop of 80 K in mantle temperature is compatible with numerical simulations of the thermal evolution of Mars[3].

The timescale of cooling is an additional point of interest and may be constrained by integration over time of the difference between internal heat production and the geochemically derived surface heat loss. Assuming reasonable values for heat-producing elements in the mantle[15] and a variation in heat flow from 38 mW m^{-2} at t_1 (the Hesperian period) to 26 mW m^{-2} at t_2 (the Amazonian period), a temperature variation may be calculated (Fig. 5). Combinations of t_1 and t_2 consistent with a temperature decrease of 80 ± 20 K imply an age for Amazonian volcanism significantly younger than 2 Gyr (Fig. 5), a result consistent with crater counts in the Tharsis region[24–26].

For the interval from the Hesperian and Amazonian periods, we calculate a mantle cooling rate of 30–40 K Gyr^{-1}, associated with a thickening of the lithosphere of 17–25 km Gyr^{-1}. This cooling rate is significantly lower than the 70–100 K Gyr^{-1} proposed for the Earth over the last 3.5 Gyr (ref. 4) The higher cooling rate for the Earth is related to a significant imbalance between heat production in the mantle and surface heat loss. The ratio of these two quantities is known as the Urey ratio; for the Earth this ratio is ~0.33 (refs 5, 6) whereas the value we derive for Mars is ~0.6–0.7. This range is close to the value of ~0.75 predicted from thermal evolution models based on Rayleigh–Nusselt scaling laws[5,6]. This comparison suggests that the low terrestrial value of the Urey ratio is the result of some specific feature of the Earth, such as a prolonged period of plate tectonics or the possible existence of a deep magma ocean that concentrates radiogenic elements[27].

Finally, we note that the mantle temperature of Mars is inferred to be barely lower than that of the Earth[4]; thus, temperature is not responsible for the lack of present volcanic activity. The decrease in the surface expression of magmatism is therefore related to thickening of the lithosphere, which terminates melting at greater depth, reducing melt volumes and the possibility of their ascending to the surface. The same may be true for other apparently extinct planetary bodies, which may remain internally hot and dynamically active. Overall, a self-consistent picture of the thermal evolution of Mars emerges, much simpler than that of the Earth, and provides a useful reference for planetary evolution.

METHODS SUMMARY

The composition in terms of major constituents (SiO_2, FeO, Al_2O_3, MgO, CaO, K_2O, Na_2O, TiO_2, Cr_2O_3, MnO and P_2O_5) of liquids produced by melting of the DW85 mantle has been determined using the thermodynamic calculator pMELTS[20] for degrees of partial melting, F, in the range of 1–20% and pressure of melting, P, in the range of 1–3 GPa. The pMELTS software has been developed primarily for terrestrial systems, but comparison with experimental data[28] allows quantitative application to more iron-rich systems relevant to Mars[10]. We then determined optimal values of P and F for each volcanic province by least-squares adjustment using pMELTS calculations.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 August 2010; accepted 1 February 2011.
Published online 6 April 2011.

1. Bibring, J.-P. *et al.* Global mineralogical and aqueous Mars history derived from OMEGA/Mars Express data. *Science* **312,** 400–404 (2006).
2. Boynton, W. V. *et al.* Concentration of H, Si, Cl, K, Fe, and Th in the low- and mid-latitude regions of Mars. *J. Geophys. Res.* **112,** E12S99 (2007).
3. Hauck, S. A. & Phillips, R. J. Thermal and crustal evolution of Mars. *J. Geophys. Res.* **107,** E75052 (2002).
4. Herzberg, C., Condie, K. & Korenaga, J. Thermal history of the Earth and its petrological expression. *Earth Planet. Sci. Lett.* **292,** 79–88 (2010).
5. Korenaga, J. Urey ratio and the structure and evolution of Earth's mantle. *Rev. Geophys.* **46,** RG2007 (2008).
6. Jaupart, C., Labrosse, S. & Mareschal, J.-C. in *Treatise on Geophysics* Vol. 7: *Mantle Dynamics* (ed. Bercovici, D.) 253–303 (Elsevier, 2007).
7. Nisbet, E. G., Cheadle, M. J., Arndt, N. T. & Bickle, M. J. Constraining the potential temperature of the Archean mantle: a review of evidence from komatiites. *Lithos* **30,** 291–307 (1993).
8. Karunatillake, S. *et al.* Chemically striking regions on Mars and Stealth revisited. *J. Geophys. Res.* **114,** E12001 (2009).
9. Gasnault, O. *et al.* Quantitative geochemical mapping of martian elemental provinces. *Icarus* **207,** 226–247 (2010).
10. El Maarry, M. R. *et al.* Gamma-ray constraints on the chemical composition of the martian surface in the Tharsis region: a signature of partial melting of the mantle? *J. Volc. Geotherm. Res.* **185,** 116–122 (2009).
11. Taylor, G. J. *et al.* Variations in K/Th on Mars. *J. Geophys. Res.* **111,** E03S06 (2006).
12. Filiberto, J. & Treiman, A. Chlorine-rich, water-poor Martian magmas. *Geochim. Cosmochim. Acta* **73** (suppl.), A376 (2009).
13. Haskin, L. A. *et al.* Water alteration and rocks and soils on Mars at the Spirit rover site in Gusev crater. *Nature* **436,** 66–69 (2005).
14. Bogard, D. D. & Johnson, P. Martian gases in an Antarctic meteorite? *Science* **221,** 651–654 (1983).
15. Dreibus, G. & Wänke, H. Mars, a volatile-rich planet. *Meteoritics* **20,** 367–381 (1985).
16. Filiberto, J., Nekvasil, H. & Lindsley, D. H. The Mars/Earth dichotomy in Mg/Si and Al/Si ratios: is it real? *Am. Mineral.* **91,** 471–474 (2006).
17. Taylor, G. J. *et al.* Bulk composition and early differentiation of Mars. *J. Geophys. Res.* **111,** E03S10 (2006).
18. Monders, A. Medard, E. & Groove, T. L. Phase equilibrium investigations of the Adirondack class basalts from the Gusev plains, Gusev crater, Mars. *Meteor. Planet. Sci.* **42,** 131–148 (2007).
19. McSween, H. Y. *et al.* Characterization and petrologic interpretation of olivine-rich basalts at Gusev Crater, Mars. *J. Geophys. Res.* **111,** E02S10 (2006).
20. Ghiorso, M. S., Hirschmann, M. M., Reiners, P. W. & Kress, V. C. III. The pMELTS: a revision of MELTS for improved calculation of phase relations and major element partitioning related to partial melting of the mantle 3 GPa. *Geochem. Geophys. Geosyst.* **3,** 1030 (2002).
21. McKenzie, D. & Bickle, M. J. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.* **29,** 625–679 (1988).
22. McGovern, P. J. *et al.* Localized gravity/topography admittance and correlation spectra on Mars: implications for regional and global evolution. *J. Geophys. Res.* **107,** E125136 (2002).
23. Belleguic, V., Lognonné, P. & Wieczorek, M. Constraints on the Martian lithosphere from gravity and topography data. *J. Geophys. Res.* **110,** E11005 (2005).
24. Neukum, G. *et al.* Recent and episodic volcanic and glacial activity on Mars revealed by the High Resolution Stereo Camera. *Nature* **432,** 971–979 (2004).
25. Hauber, E., Bleacher, J., Gwinner, K., Williams, D. & Greeley, R. The topography of low shields and associated landforms of plains volcanism in the Tharsis region of Mars. *J. Volcan. Geotherm. Res.* **185,** 1–2, 69–95 (2009).
26. Mangold, N. *et al.* Mineralogy of recent volcanic plains in the Tharsis region, Mars, and implications for platy-ridged flow composition. *Earth Planet. Sci. Lett.* **194,** 3–4,–440–450 (2009).
27. Labrosse, S., Hernlund, J. W. & Coltice, N. A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* **450,** 866–869 (2007).
28. Bertka, C. M. & Holloway, J. R. Anhydrous partial melting of an iron-rich mantle II: primary melt compositions at 15 kbar. *Contrib. Mineral. Petrol.* **115,** 323–338 (1994).
29. Mocquet, A. Rosenblatt, P. Dehant, V. & Verhoeven, O. The deep interior of Venus, Mars, and the Earth: a brief review and the need for planetary surface-based measurements. *Planet. Space Sci.* (in the press).
30. Zuber, M. T. The crust and mantle of Mars. *Nature* **412,** 220–227 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements H. Y. McSween Jr is thanked for his formal review. This work was financially supported by the Programme National de Planétologie of INSU-CNRS.

Author Contributions D.B. and M.J.T. performed calculations of melt–solid equilibria, M.M. and D.B. performed calculations of thermal evolution and O.G. provided expertise concerning the GRS data. All authors participated in the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B. (baratoux@dtp.obs-mip.fr).

METHODS

Pressure and degree of partial melting from SiO_2, FeO and Th. SiO_2 and FeO concentrations of primary melts formed in the martian mantle are functions of the degree of partial melting, *F*, and the pressure, *P*, at which the liquid was last in equilibrium with that mantle. pMELTS[20] simulations have been performed to model partial melting of the composition proposed by Dreibus and Wänke[15] (hereinafter DW85) for pressures in the range of 1–3 GPa (in steps of 0.1 GPa) and degrees of partial melting from 1 to 50%. In this way, variations of each oxide component (for example FeO) can be mapped in *P*–*F* space. With these reference grids, the values of FeO and SiO_2 determined from the GRS data may be converted to optimal values of *P* and *F* through a least-squares adjustment. Although the principle is simple, the following illustrates to what extent this inverse problem is well constrained. Considering a single oxide (SiO_2 or FeO), a family of solutions is obtained and can be represented as a curve in *P*–*F* space. The optimum *P* and *F* values are found at the intersection of the two curves (see Supplementary Information for the representation of such curves for three Hesperian and three Amazonian volcanic provinces). FeO and SiO_2 contents in the melt are both sensitive to the depth of partial melting and *P* seems to be well constrained. A larger range of values of *F* may generally explain the SiO_2 and FeO abundances within the error. Using the relationship among the ratio of thorium abundances in the mantle, that in the volcanic rocks and the degree of partial melting, the scale for the degree of partial melting can be converted into a scale indicating the thorium abundance in the mantle source. This exercise shows, for the case of Hesperian volcanism, that the range of *F* values determined from SiO_2 and FeO is compatible with the thorium abundance of the primitive martian mantle of DW85. However, for Amazonian volcanic provinces, this exercise implies that the mantle source may be depleted in thorium relative to DW85, although alternative explanations are considered below.

The influence of crystallization/assimilation on the primary melt assumption. Assuming that the surfaces described here consist of primary mantle melts leads to a self-consistent geochemical picture, at least for Hesperian volcanism. Despite this success, we have also explored the effects of fractional crystallization and assimilation of cumulates occurring before emplacement at the surface. The aim of this exercise is to assess to what extent the results of our modelling are sensitive to these processes and to determine whether the apparent variation in the thorium content of the Amazonian mantle may be explained in an alternative way. Olivine being the liquidus phase, determining the composition of the liquid as a function of the degree of fractional crystallization requires the calculation of the composition of olivine in equilibrium with the liquid, given by its Mg#. This ratio is calculated from magnesium–iron partition coefficients[31]. A third dimension has been added to the *P*–*F* grid discussed above, such that for a given GRS datum, values of *P* and *F* may be calculated for a given degree of fractional crystallization. The results of this exercise are also illustrated in Supplementary Information. For these calculations, the degree of partial melting has been fixed by the thorium content, assuming that the mantle has the thorium abundance proposed by DW85. We find that for Hesperian volcanism, geochemically acceptable values of pressure vary by less than 20% for the studied range of fractional crystallizations. For Amazonian volcanism, we find that fractional crystallization is not capable of accounting for the discrepancy between pressures determined from SiO_2 and FeO.

Similarly, the effect of assimilation of cumulates has been explored, assuming dissolution of olivine. This calculation is essentially the same as that for fractional crystallization, except that olivine is added rather than subtracted from the primary liquid. Pressure is again calculated from SiO_2 and FeO abundances, assuming the mantle thorium value of DW85 (see Supplementary Information for an illustration of the results). For Hesperian volcanism, consideration of assimilation generally leads to increasing discrepancies between the respective pressures of melting derived from SiO_2 and FeO. For the case of young volcanism, assimilation leads to a convergence of the pressures derived from SiO_2 and FeO, but values of the degree of assimilation required for overlap are considered too high to be a plausible alternative to variation in the thorium content of the mantle.

The influence of variable Mg# and/or Th of the mantle source. Given that the effects of fractional crystallization and assimilation of cumulates are not able to reconcile the geochemical data with a single *P*–*F* solution for young volcanism, a chemical modified mantle source seems inevitable for this group. One possibility is variations in Mg#. This has been explored by calculating *P*–*F* grids of SiO_2 and FeO for mantle compositions identical to that of DW85, but with modification of Mg/Fe. In an analogous way to the case of fractional crystallization, these grids have been used to derive values of pressure from SiO_2 and FeO, assuming the thorium abundance of DW85 to constrain *F*. As illustrated in Supplementary Information, alternative values of Mg# are not compatible with GRS data for Hesperian volcanism, providing further independent evidence in favour of the idea that the martian mantle is significantly richer in iron than is that of the Earth. For the younger volcanism, we note that a mantle with a Mg# greater than that proposed by DW85 is capable of providing solutions consistent with the SiO_2, FeO and thorium concentrations of surface lavas. However, given that the older Hesperian lavas are directly consistent with the DW85 composition, it would seem reasonable to assume that the Amazonian mantle is derived from that composition. In this case, episodes of melt extraction will affect both the Mg# and the thorium abundance in ways controlled by the partitioning behaviour of these elements. Two factors influence the variations in the Mg# and the thorium content of the residual mantle following melt extraction. The first is the degree of melting and the second is the fraction of the mantle that experiences this melting. Consideration of possible trajectories in Mg#–Th space (Supplementary Information) shows that whereas thorium content may be significantly affected even by small degrees of melting, variations in Mg# are relatively limited. Therefore, a possible evolution of the Mg# of the martian mantle is not ruled out, but the depletion in thorium is considered to be the dominant factor affecting the mantle source of young volcanism.

31. Toplis, M. J. The thermodynamics of iron and magnesium partitioning between olivine and liquid: criteria for assessing and predicting equilibrium in natural and experimental systems. *Contrib. Mineral. Petrol.* **149,** 22–39 (2005).

LETTER

doi:10.1038/nature09905

Metabolic trade-offs and the maintenance of the fittest and the flattest

Robert E. Beardmore[1]†*, Ivana Gudelj[1]†*, David A. Lipson[2] & Laurence D. Hurst[3]

How is diversity maintained? Environmental heterogeneity is considered to be important[1], yet diversity in seemingly homogeneous environments is nonetheless observed[2]. This, it is assumed, must either be owing to weak selection, mutational input or a fitness advantage to genotypes when rare[1]. Here we demonstrate the possibility of a new general mechanism of stable diversity maintenance, one that stems from metabolic and physiological trade-offs[3]. The model requires that such trade-offs translate into a fitness landscape in which the most fit has unfit near-mutational neighbours, and a lower fitness peak also exists that is more mutationally robust. The 'survival of the fittest' applies at low mutation rates, giving way to 'survival of the flattest'[4–6] at high mutation rates. However, as a consequence of quasispecies-level negative frequency-dependent selection and differences in mutational robustness we observe a transition zone in which both fittest and flattest coexist. Although diversity maintenance is possible for simple organisms in simple environments, the more trade-offs there are, the wider the maintenance zone becomes. The principle may be applied to lineages within a species or species within a community, potentially explaining why competitive exclusion need not be observed in homogeneous environments. This principle predicts the enigmatic richness of metabolic strategies in clonal bacteria[7] and questions the safety of lethal mutagenesis[8,9] as an antimicrobial treatment.

The enigmatic recovery of many ecotypes—defined[2] as clusters with respect to global regulation, metabolic strategies, surface properties and nutrient permeability pathways—found in a single-resource chemostat despite an absence of ecological or periodic selection[2], led us to analyse a theoretical, n-type chemostat within the framework of mutation–selection models. With evidence that within-chemostat diversity may not be maintained by frequency-dependent selection[2] and that types differ in fitness[7], it is classically supposed[10–12] that diversity must be transitory. Seeking a different explanation, we asked what conditions permit the stable maintenance of a diversity of simple organisms in a simple environment.

Consider n phenotypically distinct heritable types with the dynamics of type frequencies cast as the following mutation–selection chemostat (MSC) model:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{f} &= \varepsilon(M-I)\mathbf{f} + \mathbf{G}(S)\mathbf{f} - \langle \mathbf{G}(S), \mathbf{f}\rangle \mathbf{f}\\ \frac{\mathrm{d}}{\mathrm{d}t}\Delta &= \Delta(-d + \langle \mathbf{G}(S), \mathbf{f}\rangle)\\ \frac{\mathrm{d}}{\mathrm{d}t}S &= d(S_0 - S) - \Delta\langle \mathbf{U}(S), \mathbf{f}\rangle\end{aligned} \quad (1)$$

Here, $\mathbf{f}(t)$ is a vector of frequencies of each type, M is a stochastic mutation matrix, I is the identity (multiplicative unit) matrix, ε is the 'mutation' rate, t is time, Δ is the total density of cells per unit volume, $\mathbf{G}(S)$ is a vector of growth rates and $\mathbf{U}(S)$ is a vector of resource uptake rates in an environment with a sole carbon source of concentration S. The latter is supplied at concentration S_0 to the chemostat and all matter is lost from the culture vessel at the washout rate d. To discount factors known to affect diversity, we assume no cross-feeding[13] and that, within a population, all types are subject to the same mutation rate.

Throughout, phenotypes of cell type j, with j ranging from 1 to n, will be a function of x where $x = j/n$. The maximal uptake rate of the limiting resource by cell type j will be $xV^{\max}$ where $V^{\max}$ is a fixed constant. We therefore term x the normalized maximal uptake rate. The uptake rate of that resource, $U(x, S)$, is a Monod function proportional to the maximal uptake rate:

$$U(x, S) = V^{\max}\frac{xS}{K(x) + S} \quad (2)$$

where K, a half-saturation constant, is constrained to the maximal uptake rate and so given as a function of x. The growth rate (that is, absolute fitness) at a given resource concentration S is denoted $G(x, S)$ and defined by:

$$G(x, S) = c(x)U(x, S) \quad (3)$$

where $c(x)$ is cell yield per unit resource (for further details see Supplementary Information sections 1–7).

Equations (2) and (3) can capture two trade-offs. First, growth rate depends on cell yield and so a rate-yield trade-off arises when $c(x)$ decreases with x. Second, the growth rate depends on a cell's affinity for the carbon source and so a rate affinity trade-off occurs when $K(x)$ increases with x. Biochemically, the rate-yield trade-off is well-described and is conjectured to be a thermodynamic necessity[14] or an evolutionarily optimal strategy[15] (Supplementary Information section 4). Given that temporal fluctuations in resource abundance and ecosystem feedback can maintain diversity, we consider only what happens when resource levels converge to equilibrium, computing equilibrium loci of equation (1) for relevant mutation rates with d and S_0 fixed.

For our first computation, we set K to be constant, thus abolishing the rate-affinity trade-off and we assume that the rate-yield trade-off is convex. The resulting model captures two recognized evolutionary principles (Supplementary Fig. 12a). First, at low mutation rates the fittest type and its mutational neighbours dominate; this is classical 'survival of the fittest' with limited diversity maintained at mutation–selection equilibrium (see ref. 14). At high mutation rates another type dominates in which the fitness landscape is flat and organisms are robust to mutation; this is 'survival of the flattest'[4–6]. The transition from survival of the fittest to survival of the flattest, as the mutation rate increases, passes through a region in which both quasispecies are abundant. This we term the principle of the maintenance of the fittest and flattest.

With just one trade-off we find a relatively narrow maintenance region. However, there potentially exist multiple trade-offs even in simple organisms (Supplementary Information section 4). Consider, for

[1]Department of Mathematics, Imperial College London, Huxley Building, 180 Queen's Gate, London SW7 2AZ, UK. [2]Department of Biology, San Diego State University, San Diego, California 92182-4614, USA. [3]Department of Biochemistry and Biology, University of Bath, Claverton Down, Bath, BA2 7AY, UK. †Present address: Biosciences, Geoffrey Pope Building, Streatham Campus, University of Exeter, Exeter, Devon, EX4 4SB, UK.
*These authors contributed equally to this work.

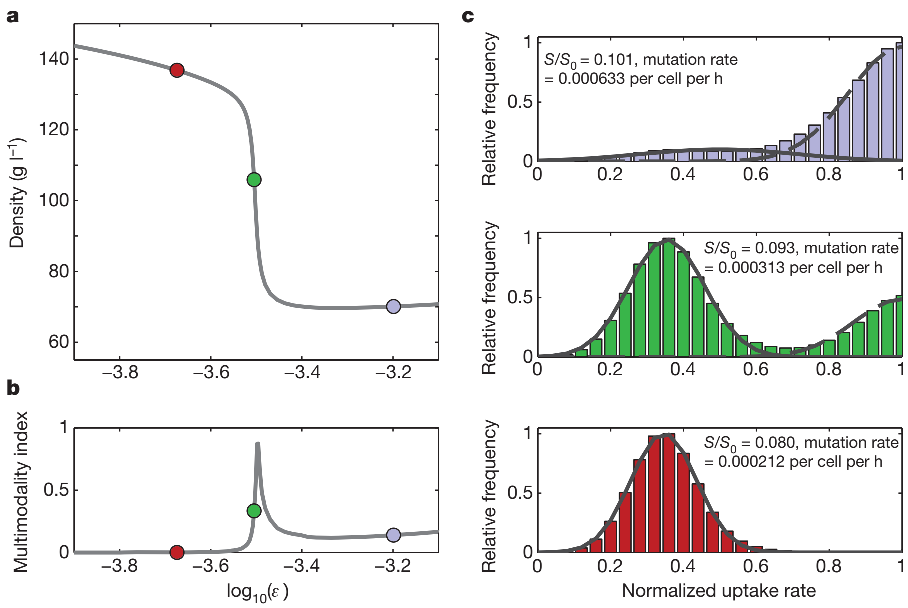


Figure 1 | Maintenance of the fittest and flattest with known parameter values. **a**, Using a parameterization of the MSC equation with rate-affinity and rate-yield trade-off data from empirical studies of *S. cerevisiae* (see Supplementary Fig. 5), the diagram shows the locus of steady-state densities for different mutation rates. **b**, The cited multimodality index is the ratio of biomass supported under each quasispecies. **c**, Normal distributions have been fitted to each quasispecies to provide a guide.

example, what happens in our model if we also permit the rate-affinity trade-off. Even supposing this trade-off to be linear, the region through which the fittest and the flattest are maintained can be broader than the simpler model (Supplementary Fig. 12b). Examination of more complex models reveals that the more parameters through which trade-offs exist, the broader the span of mutation rates through which diversity is maintained (Supplementary Information section 6.4, compare with ref. 16). A quantitative model incorporating empirical rate-yield and rate-affinity data captures the maintenance of the fittest and the flattest (Fig. 1, and Supplementary Information section 6.4.1). The mutation

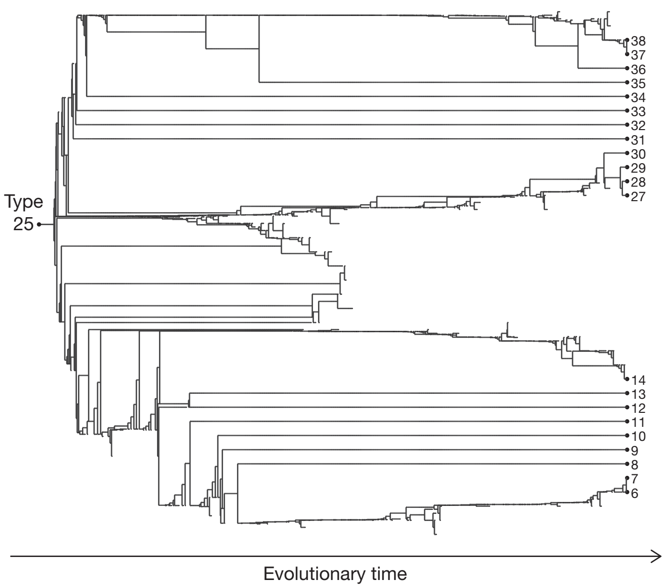


Figure 2 | The maintenance of the fittest and the flattest is not a simple mutation–selection equilibrium. A single type (number 25) seeded a clonal population subject to rate-yield and rate-affinity trade-offs that diverged into two lineages of quasispecies of efficient and inefficient generalists. For the trade-offs see Supplementary Fig. 17; for further details see Supplementary Information section 7.3.

rate required for co-maintenance appears to be realistic (Supplementary Information section 6.4.2). We also modelled populations using individual-based stochastic simulations and the results are unaffected (Methods and Supplementary Information section 7). We conclude that the region of maintenance is broad enough to be relevant even to metabolically simple organisms and is likely to be especially relevant in metabolically rich microorganisms.

The trade-offs assumed in our model have been observed within species of microorganisms[17–19] and within communities[20] (Supplementary Information section 4). In the simplest model with a single trade-off, the trade-off must have a convex component to enable co-maintenance. A convex component in the rate-yield trade-off is possible for pathways with opposing properties in rate and yield[17] and has been observed elsewhere in a eukaryote (yeast[18]; Supplementary Fig. 5), a prokaryote (*Escherichia coli*; Supplementary Fig. 4) and in a community (soil microbes[20]; Supplementary Fig. 3). A rate-affinity trade-off with a convex component is also observed in yeast (Supplementary Fig. 5). With multiple trade-offs, their rich variety can provide sufficiently complex fitness landscapes for the principle to work (Supplementary Information section 6.4). Although here we consider well-described metabolic trade-offs, the same principle may apply to other metabolic or physiological trade-offs. For example the balance between stress protection and nutritional competence can also support the maintenance of diversity (S. Nilsson, T. Ferenci, K. Haynes and I.G., unpublished work).

Prior explanations for the maintenance of diversity in homogeneous, well-mixed environments (weak selection/neutrality, mutation–selection equilibrium and genotypic negative frequency-dependence) do not account for these findings. If we consider a population with a broad diversity of possible types and let it evolve to steady state but without further mutations, one type, the fittest, predominates (Supplementary Information section 7.2). The various types are therefore not maintained by any weakness of selection.

To exclude mutation–selection equilibrium we tracked lineages using individual-based simulations, asking whether mutants derived from one quasispecies can contribute biomass to another. Importantly, we find that the two quasispecies can be mutationally independent (Fig. 2). If we force a mutational barrier between the two peaks after formation, their stable maintenance remains and may be enhanced

(Supplementary Information section 7.4). The types are therefore not maintained by mutational input from one quasispecies alone and hence simple mutation–selection equilibrium is not a full explanation. Similarly, introducing two mutationally distinct populations with no recent common ancestor (that is, different species) competing for the same resource can also result in maintenance, demonstrating relevance to diversity between species in a community.

The maintenance of diversity is not owing to classical frequency-dependent selection. If it were, the diagnostic of enforcing competition between two types, one from each quasispecies and in the absence of mutation, would result in a stable balance between the two[12]. However, in the absence of mutation here, one fittest type excludes all others (Supplementary Fig. 19a). Considering just two types from each fitness peak in competition, the fittest type from the highest peak excludes the other (Supplementary Fig. 19a and b).

Although our principle is logically distinct from prior explanations, it does bear some resemblance to these. Notably, the co-maintenance of quasispecies can be configured as a form of quasispecies-level (not genotype level) negative frequency dependence (Fig. 3a and b). If we perturb the frequencies of the quasispecies, those below their equilibrium value increase in frequency while those above decrease in frequency (Fig. 3b), a property that requires mutational connectivity within quasispecies but not between quasispecies (Supplementary Information section 7.4, and Supplementary Fig. 23). This dynamic is mediated by shifts in resource

BOX 1

A toy model explains how two quasispecies may be co-maintained

Consider two clusters of types forming quasispecies 'a' and 'A'. The fitness of quasispecies a is G_a, the fitness of quasispecies A is G_A and the relative frequency of the totality of all types in quasispecies A is f_A. We assume that negative frequency dependence is exhibited (see Fig. 3, and Supplementary Information section 8):

$$\frac{G_A}{G_a} = r_0 - f_A(r_0 - r_1) \tag{4}$$

where $r_0 > r_1 > 1$ are fixed constants and $0 \le f_A \le 1$. This implies $G_A > G_a$, meaning A is the fitter of the two quasispecies. We now assume three types of mutational event:

ME1: a→a or A→A: intra-quasispecies mutation
ME2: a→A or A→a: inter-quasispecies mutation
ME3: a→ dead or A→ dead: lethal mutations

To discount mutation–selection equilibrium, we assume an absence of inter-quasispecies mutations (ME2) and, for simplicity, the fitness effects of intra-quasispecies mutational events (ME1) are ignored.

Now, suppose that the loss of cell types from quasispecies a and A owing to death (ME3) reduces the overall fitness of a and A, so that:

$$w_a = G_a - \mu_a \text{ and } w_A = G_A - \mu_A \tag{5}$$

where $\mu_a > 0$ and $\mu_A > 0$. Thus w_a and w_A are the quasispecies fitnesses in the presence of mutation. Importantly, because *a* is the more mutationally robust (flatter) quasispecies, quasispecies *a* has a lower rate of lethal mutation, so we assume $\mu_A > \mu_a$ and set $\alpha = \mu_a/\mu_A$ with $0 < \alpha < 1$.

This toy model is consistent with survival of the fittest and the flattest, respectively. Survival of the fittest is relevant when $w_A > w_a$ which occurs when $G_a(r_1 - 1)/(1 - \alpha) > \mu_A$. Survival of the flattest is relevant when $w_A < w_a$, that is, when $G_a(r_0 - 1)/(1 - \alpha) < \mu_A$. Co-maintenance of both quasispecies is possible when $w_A = w_a$; this occurs when mutation rates take on intermediate values:

$$\frac{G_a(r_1 - 1)}{1 - \alpha} < \mu_A < \frac{G_a(r_0 - 1)}{1 - \alpha}$$

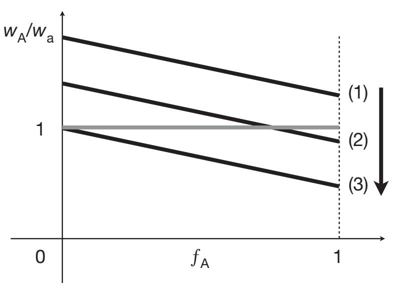
(see Box figure). This principle requires $G_A > G_a$, $\mu_A > \mu_a$ and that negative frequency-dependent selection acts between the two quasispecies. Were we to assume that mutations in quasispecies *a* had the same probability of being lethal as those in quasispecies A, (that is, $\mu_A = \mu_a$), A would be the fitter for all mutation rates. Similarly, in the absence of negative frequency-dependence, $r_0 = r_1$, equation (4) would become:

$$G_A = r_0 G_a \tag{6}$$

Allowing for $\mu_a = \alpha\mu_A$ with $0 < \alpha < 1$ and substituting equation (6) into equation (5), we obtain:

$$\frac{w_A}{w_a} = \frac{G_a r_0 - \mu_A}{G_a - \alpha\mu_A} \tag{7}$$

In equation (7) there is just one value of μ_A for which a non-robust form of co-maintenance may occur, namely $\mu_A = G_a(r_0 - 1)/(1 - \alpha)$ (Supplementary Fig. 26b).



Box figure | The relative fitness of the fit quasispecies A as a function of its frequency f_A. For increasing mutation rates, given by the direction of the arrow, the survival of the fittest quasispecies ((1) gives way to survival of the flattest (3) via a regime of co-maintenance (2)).

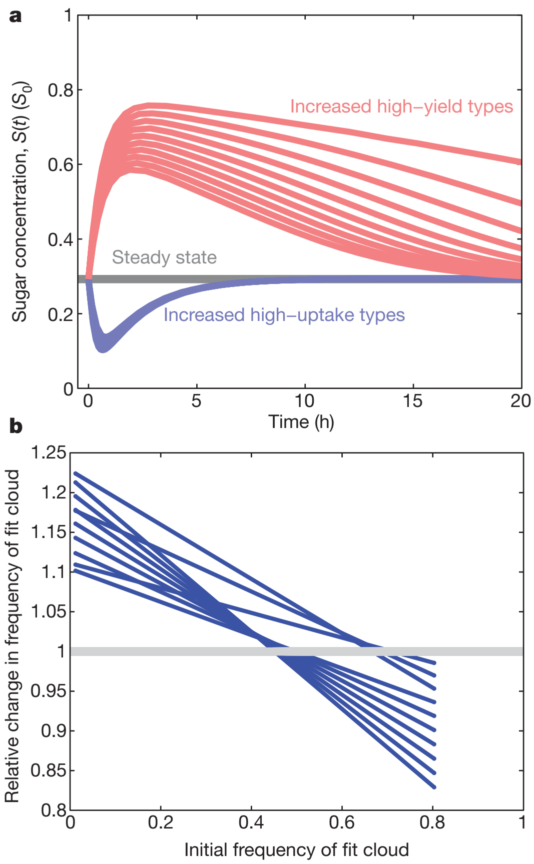


Figure 3 | Quasispecies negative frequency-dependent selection. **a**, Sugar dynamics after altering cloud frequencies. The MSC equation maintains distinct quasispecies though negative frequency-dependent selection mediated by the abiotic environment. Shown are typical responses in resource concentration resulting from changes in the frequency of fit quasispecies (high uptake) and flat quasispecies (high yield). Removing 'fit cells' momentarily decreases sugar concentration; removing 'flat cells' increases it. **b**, This results in negative frequency dependence at the level of quasispecies. Ten tests of the MSC model for negative frequency dependence are shown. For trade-offs see Supplementary Fig. 8b. The grey line indicates the line of equal fitness.

levels that arise upon perturbation: quasispecies differ in the rate at which they consume resources, and raising the frequency of either one changes those resource levels (Fig. 3a, and Supplementary Information section 8). Owing to the trade-offs, this in turn affects relative growth rates differently, thus ensuring that the quasispecies and resource levels return to their prior equilibrium (Fig. 3a). For example, when the frequency of high-rate 'fit' types is increased above equilibrium, the resource concentration reduces and the concomitant increased growth of the high-efficiency 'flatter' cells eventually resets resource levels.

The principle of the maintenance of the fittest and the flattest requires more than some form of negative-frequency dependence. A toy population-genetical formulation highlights a set of minimal requirements (Box 1). Stable diversity is seen in the absence of spatial or temporal heterogeneity and requires neither direct biotic interactions nor mutation-rate heterogeneity. For the principle to hold, as seen in the toy model, there needs to be (1) negative frequency-dependent selection between the quasispecies (arising here through resource competition and metabolic/physiological trade-offs), (2) the quasispecies that is less fit must be more robust to mutations (that is, flatter) and (3) the type of a parent and its offspring must be positively correlated to ensure continuity of the quasispecies. The same principle can be extended to the maintenance of more than two quasispecies.

Our model is consistent with the unexpected outcome of chemostat experiments. Close examination of the metabolic efficiencies of chemostat-derived bacteria found remarkable metabolic diversity, some with high rates, some with high yield and many rate–yield hybrids[7], as our model predicts. This is contrary to the expectations of alternative models[14].

The principle is not simply a new theoretical formulation to explain diversity found in microbial communities—the results apply in principle to any clonally derived population, such as tumours and their metastases. Although tumour cells are typically clonal derivates, they are often highly heterogeneous[21]. This may reflect niche differentiation[21]. However, much as with chemostat bacterial populations, cancer cells often show paradoxically[22] inefficient metabolism[23], notably low ATP-yield aerobic glycolysis[23]; our model suggests why inefficient metabolism might be found. The observed metabolic heterogeneity in response to chemotherapy[24] is thus explicable.

We assumed that all types have the same mutation rate. This need not, however, be a realistic assumption even within species, not least because bacteria can increase their mutation rate when subject to stress[25]. How higher rates of mutation within a population affect diversity and density is relevant to understanding anti-microbial treatments that increase mutation rate (lethal mutagenesis[8]). Importantly, the survival of the fittest and flattest questions the safety of this therapy. Whereas recent models[26] suggest that pathogen densities should decrease linearly as mutation rates increase, our work suggests that increasing mutation rates can increase both density and diversity by pushing the population to a slightly lower but much flatter fitness peak (Supplementary Fig. 11).

METHODS SUMMARY

Stochastic simulations. An individual-based stochastic model was constructed consisting of n heritable types, each with a unique phenotype consisting of cell yield, resource affinity and maximal uptake rate. Each heritable type j was numbered from 1 to n (we note that a clonal population of n types is equivalent to having n alleles at one locus) and the phenotypic state x_j of type j is then $x_j = j/n$ and all phenotypes are derived from this number. The environment contains resource molecules that can be adsorbed by each cell and symmetric division occurs only when a given number of molecules is metabolized, a number that defines the efficiency of each cell. Mutations are introduced with a fixed probability of occurrence per unit time that convert a cell of type j to either $j+1$ or $j-1$; the latter is not a necessary assumption but is conservative and prevents peak-jumping through the generation of 'hopeful monsters' (Supplementary Information sections 6.5 and 7.5). Relaxation of this assumption can even promote diversity (Supplementary Information section 7.5).

Analysis of the MSC model. With mutation rate as the bifurcation parameter, equilibrium states that branch transcritically from the trivial solution of equation (1) (where $\Delta = 0$) were computed using the pseudo arc-length computational strategy[27] (with LU decomposition for the Newton solver, 10^{-14} residual, implemented in Matlab version 7.9, 2009b) and extended into the region of lowest mutation rates. The global stability of the unique equilibrium of equation (1) can be proved under restrictions; where proofs are not available eigenvalue computations and the geometric structure (bifurcations) of solution branches was used to deduce local stability.

Received 1 December 2010; accepted 7 February 2011.
Published online 27 March 2011.

1. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31,** 343–366 (2000).
2. Maharjan, R., Seeto, S., Notley-McRobb, L. & Ferenci, T. Clonal adaptive radiation in a constant environment. *Science* **313,** 514–517 (2006).
3. Gudelj, I., Beardmore, R. E., Arkin, S. S. & MacLean, R. C. Constraints on microbial metabolism drive evolutionary diversification in homogeneous environments. *J. Evol. Biol.* **20,** 1882–1889 (2007).
4. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E. & Adami, C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* **412,** 331–333 (2001).
5. Sanjuán, R., Cuevas, J. M., Furio, V., Holmes, E. C. & Moya, A. Selection for robustness in mutagenized RNA viruses. *PLoS Genet.* **3** (6), e93, 939–946 (2007).
6. Codoñer, F. M., Daros, J. A., Sole, R. V. & Elena, S. F. The fittest versus the flattest: Experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS Pathog.* **2** (12), e136, 1187–1193 (2006).
7. Maharjan, R. P., Seeto, S. & Ferenci, T. Divergence and redundancy of transport and metabolic rate-yield strategies in a single *Escherichia coli* population. *J. Bacteriol.* **189,** 2350–2358 (2007).

8. Loeb, L. A. *et al.* Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc. Natl Acad. Sci. USA* **96,** 1492–1497 (1999).
9. Bull, J. J. & Wilke, C. O. Lethal mutagenesis of bacteria. *Genetics* **180,** 1061–1070 (2008).
10. Koch, A. L. The pertinence of the periodic selection phenomenon to prokaryote evolution. *Genetics* **77,** 127–142 (1974).
11. Levin, B. R. Periodic selection, infectious gene exchange and the genetic structure of *E. coli* populations. *Genetics* **99,** 1–23 (1981).
12. Turner, P. E., Souza, V. & Lenski, R. E. Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology* **77,** 2119–2129 (1996).
13. Pfeiffer, T. & Bonhoeffer, S. Evolution of cross-feeding in microbial populations. *Am. Nat.* **163,** E126–E135 (2004).
14. Pfeiffer, T., Schuster, S. & Bonhoeffer, S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292,** 504–507 (2001).
15. Molenaar, D., van Berlo, R., de Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5,** 323, doi:10.1038/msb.2009.82 (2009).
16. Doebeli, M. & Ispolatov, I. Complexity and diversity. *Science* **328,** 494–497 (2010).
17. Novak, M., Pfeiffer, T., Lenski, R. E., Sauer, U. & Bonhoeffer, S. Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*. *Am. Nat.* **168,** 242–251 (2006).
18. Weusthuis, R. A., Pronk, J. T., van den Broek, P. J. & van Dijken, J. P. Chemostat cultivation as a tool for studies on sugar transport in yeasts. *Microbiol. Mol. Biol. Rev.* **58,** 616–630 (1994).
19. Kreft, J. U. Biofilms promote altruism. *Microbiology* **150,** 2751–2760 (2004).
20. Lipson, D. A., Monson, R. K., Schmidt, S. K. & Weintraub, M. N. The trade-off between growth rate and yield in microbial communities and the consequences for under-snow soil respiration in a high elevation coniferous forest. *Biogeochemistry* **95,** 23–35 (2009).
21. Marusyk, A. & Polyak, K. Tumor heterogeneity: Causes and consequences. *Biochim. Biophys. Acta. Rev. Cancer* **1805,** 105–117 (2010).
22. Gillies, R. & Gatenby, R. Adaptive landscapes and emergent phenotypes: why do cancers have high glycolysis? *J. Bioenerg. Biomembr.* **39,** 251–257 (2007).
23. Warburg, O. On the origin of cancer cells. *Science* **123,** 309–314 (1956).
24. Krak, N. *et al.* Blood flow and glucose metabolism in stage IV breast cancer: heterogeneity of response during chemotherapy. *Mol. Imaging Biol.* **10,** 356–363 (2008).
25. Bjedov, I. *et al.* Stress-induced mutagenesis in bacteria. *Science* **300,** 1404–1409 (2003).
26. Martin, G. & Gandon, S. Lethal mutagenesis and evolutionary epidemiology. *Phil. Trans. R. Soc. Lond. B* **365,** 1953–1963 (2010).
27. Keller, H. B. in *Applications of Bifurcation Theory* (ed. Rabinowitz, P.) (Academic Press, 1977).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Burch, M. Doebeli and T. Ferenci for discussions. L.D.H. is a Royal Society Wolfson Research Merit Award Holder, R.E.B. holds an EPSRC Leadership Fellowship, and I.G. holds a NERC Advanced Research Fellowship.

Author Contributions R.E.B. and I.G. wrote the paper, conceived the paper, designed analyses and performed analysis, D.A.L. wrote the paper and performed analysis, L.D.H. wrote the paper, conceived the paper and designed analyses.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.D.H. (l.d.hurst@bath.ac.uk).

doi:10.1038/nature09972

A transient placental source of serotonin for the fetal forebrain

Alexandre Bonnin[1,2], Nick Goeden[1], Kevin Chen[3], Melissa L. Wilson[4], Jennifer King[4], Jean C. Shih[3], Randy D. Blakely[5,6], Evan S. Deneris[7] & Pat Levitt[1,2]

Serotonin (5-hydroxytryptamine or 5-HT) is thought to regulate neurodevelopmental processes through maternal–fetal interactions that have long-term mental health implications. It is thought that beyond fetal 5-HT neurons there are significant maternal contributions to fetal 5-HT during pregnancy[1,2] but this has not been tested empirically. To examine putative central and peripheral sources of embryonic brain 5-HT, we used *Pet1*$^{-/-}$ (also called Fev) mice in which most dorsal raphe neurons lack 5-HT[3]. We detected previously unknown differences in accumulation of 5-HT between the forebrain and hindbrain during early and late fetal stages, through an exogenous source of 5-HT which is not of maternal origin. Using additional genetic strategies, a new technology for studying placental biology *ex vivo* and direct manipulation of placental neosynthesis, we investigated the nature of this exogenous source. We uncovered a placental 5-HT synthetic pathway from a maternal tryptophan precursor in both mice and humans. This study reveals a new, direct role for placental metabolic pathways in modulating fetal brain development and indicates that maternal–placental–fetal interactions could underlie the pronounced impact of 5-HT on long-lasting mental health outcomes.

Fetal 5-HT dysfunction is implicated in developmental programming by altering brain circuit formation[4], which later translates into abnormal adult behaviours[5,6]. In humans, risk alleles in genes involved in 5-HT function combine with early adverse experiences during development to affect adult-onset mental illnesses[7]. Furthermore, polymorphisms in the 5-HT transporter (*SLC6A4*, which encodes SERT) and 5-HT receptors, which are expressed early in brain development[8], are associated with neurodevelopmental disorders such as autism spectrum disorder and schizophrenia[9,10]. A puzzling issue regarding the role of 5-HT in fetal brain development is that receptors, transporters and degrading enzymes for 5-HT often appear before the development of 5-HT innervation[11] itself, suggesting the existence of an exogenous source of 5-HT at early stages of development. The most biologically influential source of exogenous 5-HT is claimed to be maternal[1,2] but there is a lack of experimental data to support such a mechanism of developmental programming.

Analysis of fetal *Pet1*$^{-/-}$ mice provides an opportunity to assess potential extra-embryonic sources of 5-HT, as only ~30% of dorsal raphe 5-HT neurons can be detected in null mice, which also express low levels of tryptophan hydrolase 2 (TPH2), aromatic l-amino acid decarboxylase (AADC; also called dopa decarboxylase, Ddc), SERT and other markers of the serotonergic phenotype[3]. To assess whether dorsal raphe neurons are also the sole source of brain 5-HT during development, we compared the concentration of 5-HT in embryonic brains collected from *Pet1*$^{-/-}$ and wild-type littermates from embryonic day (E) 10.5, the onset of 5-HT synthesis in dorsal raphe neurons, to E17.5, when 5-HT axons are fully deployed throughout the forebrain. High-performance liquid chromatography (HPLC) was used to measure the concentration of 5-HT in the mid/hindbrain region (termed 'hindbrain'), which contains 5-HT cell bodies and proximal axons, and in the forebrain, which contains only distal 5-HT axons[12,13]. Consistent with dorsal raphe neurons providing the main source of 5-HT in the hindbrain, the 5-HT concentration was lower in *Pet1*$^{-/-}$ hindbrains than in those of wild-type mice at every age tested (Fig. 1a). Notably, in the *Pet1*$^{-/-}$ forebrain, 5-HT levels were statistically indistinguishable from wild type at E10.5 to E15.5; however, large differences emerged at E16.5 (Fig. 1b). This is consistent with dorsal raphe axons being the major source of forebrain 5-HT at this and later ages but not earlier (Supplementary Fig. 2 and ref. 3). Remarkably, even before the arrival of 5-HT axons in the ventral forebrain (E10.5–E12.5), low levels of 5-HT were detected (Fig. 1b). Normally, over the next three embryonic days, progressively more 5-HT axons grow into the forebrain[13] (Fig. 1c–e). In the *Pet1*$^{-/-}$ forebrain, however, there is a marked reduction in 5-HT axon density compared to wild type (Fig. 1c–i), even though total tissue 5-HT concentrations are comparable. The density and distribution of thalamocortical axons, which also express SERT and can take up 5-HT[14], are similar in *Pet1*$^{-/-}$ and wild-type mice (Fig. 1c, f). These results reveal a complex regulation of 5-HT in the fetal brain, with dorsal raphe serotonergic neurons and axons representing the major source of 5-HT in the hindbrain and at later embryonic stages in the forebrain but not the main source of 5-HT in the early developing forebrain.

The greater decrease in total tissue 5-HT concentration in the hindbrain than in the forebrain in *Pet1*$^{-/-}$ mice suggests a differential contribution of non-dorsal-raphe sources in these regions. Alternatively, because 5-HT degradation enzyme (monoamine oxydase A; MAO-A) activity is higher in the hindbrain than in the forebrain at early stages of development[15], a differential degradation of 5-HT across the two brain regions may account for the difference. Consistent with this latter possibility, 5-hydroxyindoleacetic acid (5-HIAA) concentration in the E14.5 s*Maoa*$^{-/-}$ mouse (which lacks MAOA enzymatic activity and cannot efficiently degrade 5-HT[16]) were decreased 3.4-fold in the forebrain, but 6.1-fold in the hindbrain compared to wild type littermates (Supplementary Fig. 3). In contrast, at E16.5, 5-HIAA concentrations were decreased to a similar extent in the s*Maoa*$^{-/-}$ forebrain and hindbrain (3.6- and 3.1-fold respectively). A downregulation of MAOA activity in the *Pet1*$^{-/-}$ forebrain before E16.5 could explain the normal concentrations of 5-HT measured in the region. To test this possibility, we quantified MAOA activity in the forebrain at E14.5 and showed that it was not different in *Pet1*$^{-/-}$, *Pet1*$^{+/-}$ and *Pet1*$^{+/+}$ forebrains (Supplementary Table 2). Furthermore, 5-HIAA concentrations were not different in the *Pet1*$^{-/-}$ and *Pet1*$^{+/+}$ forebrains at E12.5 and E14.5 (Supplementary Fig. 3c). The data demonstrate that MAO-A activity is not downregulated in the *Pet1*$^{-/-}$ forebrain.

The non-dorsal-raphe origin of 5-HT in the early fetal forebrain could arise from several sources. Because adult forebrain catecholaminergic

[1]Zilkha Neurogenetic Institute, Keck School of Medicine of USC, Los Angeles, California 90089, USA. [2]Department of Cell & Neurobiology, Keck School of Medicine of USC, Los Angeles, California 90089, USA. [3]Department of Pharmacology & Pharmaceutical Sciences, USC School of Pharmacy, Los Angeles, California 90089, USA. [4]Departments of OB/GYN & Preventive Medicine, Keck School of Medicine of USC, Los Angeles, California 90089, USA. [5]Department of Pharmacology, Vanderbilt University Medical Center, Nashville, Tennessee 37221, USA. [6]Center for Molecular Neuroscience, Vanderbilt University Medical Center, Nashville, Tennessee 37221, USA. [7]Department of Neuroscience, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, USA.

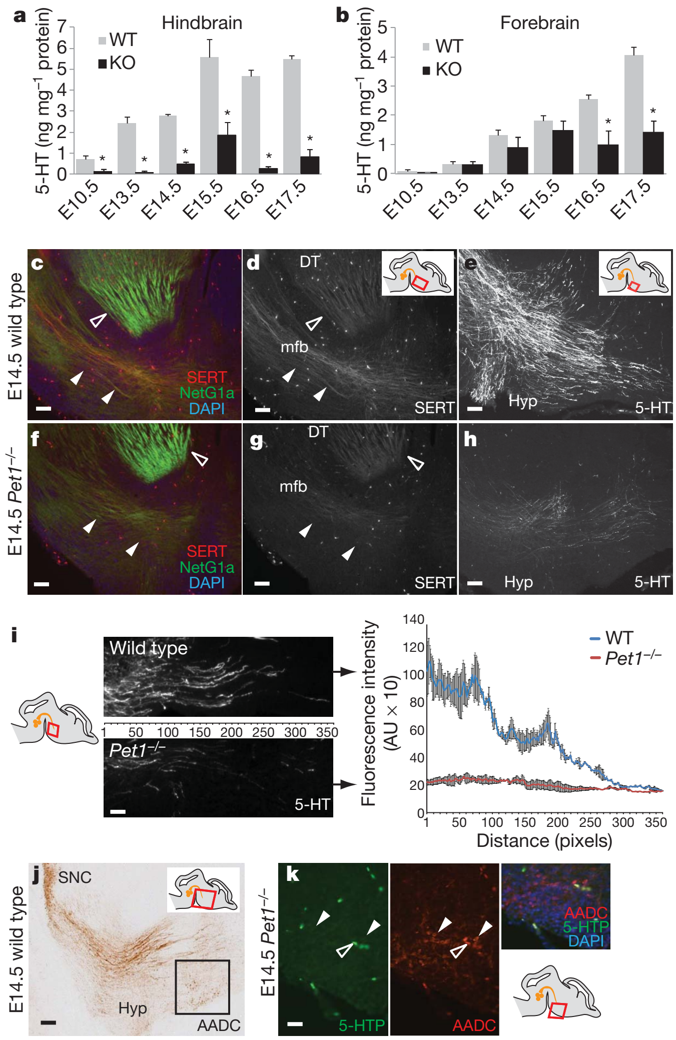


Figure 1 | Comparison of fetal 5-HT concentrations in the hindbrain and forebrain of *Pet1*$^{-/-}$ and wild type embryos from E10.5 to E17.5. a, 5-HT concentration in the *Pet1*$^{-/-}$ (KO) hindbrain is significantly lower than in wild-type (WT) littermates at every age tested. **b**, In contrast, 5-HT concentrations in the *Pet1*$^{-/-}$ forebrain are not significantly different from wild-type littermates from E10.5 to E15.5 but become significantly lower at E16.5 and E17.5 ($n \geq 6$ embryos per genotype per age; *, $P < 0.005$; one-way analysis of variance (ANOVA); data are presented as means ± s.e.m.). **c**–**h**, Serotonergic axons (SERT^{+}) and dorsal thalamic (DT) axons (NetG1a^{+}) immunostained on sagittal sections at E14.5 in wild-type (**c**) and *Pet1*$^{-/-}$ (**f**) embryos (regions shown correspond to the red box in the drawings). In wild-type E14.5 embryos (**d**) SERT^{+} axons grow ventrally into the forebrain through the medial forebrain bundle (mfb, white arrowheads); SERT also labels dorsal thalamic axons at this age (open arrowhead). In comparison, very few SERT^{+} axons remain in the *Pet1*$^{-/-}$ (**g**). The pattern and density of SERT^{+} DT axons appear to be unaffected in the *Pet1*$^{-/-}$ forebrain. Scale bars: 50 µm. The rostral-most extent of ingrowing serotonergic axons immunolabelled with 5-HT in the wild type (**e**) shows numerous 5-HT^{+} axons, some of which diverge towards the hypothalamus (Hyp). In contrast, only few 5-HT^{+} axons remain in the *Pet1*$^{-/-}$ forebrain (**h**). Scale bars: 20 µm. **i**, Densitometric analysis of 5-HT^{+} axons in the most rostral part of the medial forebrain bundle (region indicated in the diagram) at E14.5 confirms fewer axons in the *Pet1*$^{-/-}$. AU, arbitrary units. **j**, AADC staining identifies dopaminergic neurons in the substantia nigra pars compacta (SNC), along with dopaminergic and serotonergic axons coursing through the ventral forebrain at E14.5, and also AADC^{+} catecholaminergic neurons present in the hypothalamus (Hyp, black box). Scale bar: 100 µm. **k**, AADC^{+} neurons in the hypothalamic region (red box in bottom right drawing; middle panel, white arrowheads) are 5-HTP-negative. Open arrowheads indicate fluorescence from blood vessels. Scale bar: 25 µm.

and dopaminergic neurons express the AADC enzyme, these cells can ectopically synthesize 5-HT, albeit after administration of large doses of the precursor 5-HTP[17]. We tested the possibility that embryonic AADC-expressing neurons could ectopically produce 5-HT in the early *Pet1*$^{-/-}$ forebrain. Consistent with measures of unaltered dopamine concentration (Supplementary Fig. 1), AADC immunostaining revealed normal catecholaminergic and dopaminergic neuron and axon density in the *Pet1*$^{-/-}$ forebrain (Fig. 1j). Furthermore, catecholaminergic neurons present in the *Pet1*$^{-/-}$ hypothalamus do not exhibit ectopic 5-HTP or 5-HT immunoreactivity (Fig. 1k), consistent with there being no local cellular source of 5-HT in the *Pet1*$^{-/-}$ forebrain.

In the developing forebrain, E10.5–E15.5 is the period of pronounced neurogenesis and axon growth. As 5-HT modulates both processes[4,10], it is essential that its availability be regulated during this time. It is remarkable, therefore, that over this time period, even in the absence of 5-HT axons, the concentration of 5-HT in the *Pet1*$^{-/-}$ forebrain is normal. Possible exogenous sources include the embryonic gut, the maternal blood through the placenta, or the placenta itself. We ruled out the embryonic gut as a source because expression of the 5-HT biosynthetic enzyme tryptophan hydroxylase (TPH1), which provides blood 5-HT, begins late (E15.5) in fetal enterochromaffin cells[1,18]. To test the possibility that maternal 5-HT is transferred to the fetal brain, we examined brains from fetuses of *Sert* knockout (*Sert*$^{-/-}$) dams; total blood and platelets in these dams contain virtually no 5-HT[19]. This absence of blood 5-HT is attributed to a failure of uptake by platelets, to rapid degradation of the remaining free plasma 5-HT in the liver and to compensatory uptake by other transporters in the gut[19]. Despite this, the concentration of 5-HT is not different in the forebrain of *Sert*$^{+/-}$ E12.5 embryos from *Sert*$^{-/-}$ or wild-type dams (Supplementary Table 1), indicating that maternal blood 5-HT is not the main source of fetal blood and forebrain 5-HT at early stages of development.

We considered the alternative hypothesis that the essential amino acid tryptophan, originating from the pregnant dam, would be converted to 5-HT in the placenta and delivered to the fetal circulation. Injection of tryptophan in pregnant dams increases 5-HT concentration in the fetal brain[20] but the precise location of the synthetic conversion has never been identified. Quantitative RT-PCR of placental tissue detected transcripts of *Tph1* and *Aadc*, but not *Pet1* (Supplementary Fig. 4a). Immunocytochemistry confirmed that TPH1 and AADC proteins are expressed in the syncytiotrophoblastic cell layer of the placenta at E10.5–E14.5 (Fig. 2c–h and Supplementary Fig. 4e–j). The placenta thus has the necessary machinery to synthesize 5-HT. We tested for placental 5-HT neosynthesis and showed that both 5-HTP and 5-HT neosyntheses occur in placenta and fetal hindbrain extracts incubated with tryptophan (Fig. 2a) as early as E10.5 (Supplementary Fig. 4b). Placental 5-HT synthesis capacity was greater at E14.5 than at E18.5, whereas the converse was true in the hindbrain, consistent with changes in *Tph1* transcript expression (data not shown). The capacity for placental 5-HT synthesis at E14.5 was not affected by the absence of embryonic *Pet1* gene expression (Supplementary Fig. 4c). This synthetic capacity is not unique to mice, as human placental fetal villi at 11 weeks of gestation showed robust 5-HT neo-synthesis (Fig. 2b), indicating that a placental source of 5-HT is important for human fetal development.

Although not providing conclusive proof, these data are consistent with the possibility that neosynthesis and transport of 5-HT from a maternal tryptophan precursor occurs in an intact placenta. We addressed this with two strategies. First, we developed a novel *ex vivo* technology for regulating placental organ perfusion, thus allowing the presentation of maternal precursor and the collection of fetal perfusate in intact, live murine placentas (Fig. 2i). Within 15 min of tryptophan injection through the maternal uterine artery, there was a large accumulation of newly synthesized 5-HT that passes through the fetal placental circulation (Fig. 2j), demonstrating that the live placenta is able to convert tryptophan to 5-HT and release the neurotransmitter into the fetal circulation. In contrast, when 5-HT (1.5 nM) was injected into the uterine artery, only 0.32 ± 0.16% of the maternal free 5-HT was transferred to the fetal umbilical vein during a 30 min perfusion

period. In a second strategy, we specifically blocked placental TPH1 enzyme activity by microinjecting small volumes of the TPH inhibitor p-chlorophenyalanine (PCPA) directly into the labyrinth zone of E14.5 placentas *in utero* (Fig. 3a). To minimize nonspecific effects due to diffusion of the drug into the maternal and fetal blood compartments, placental and fetal brain tissues were collected after a short 30-min period of drug exposure. This pharmacological manipulation reduced 5-HT levels in the placenta but did not reduce 5-HT in the fetal hindbrain, indicating that exposure to PCPA was too short to inhibit TPH2 activity in dorsal raphe serotonergic neurons (Fig. 3b). Nevertheless, the brief exposure to PCPA resulted in a significant decrease in fetal forebrain 5-HT levels (Fig. 3b). The data demonstrate that an exogenous source of 5-HT produced in the placenta is required to maintain normal levels of forebrain 5-HT during early stages of forebrain development (Supplementary Fig. 5).

The concept that 5-HT from maternal blood could be transferred to the fetal circulation after crossing the placenta is widely accepted[1,2] but direct or indirect transfer of the molecule has never been demonstrated. Although uptake of exogenous 5-HT by syncytiotrophoblasts in the human and mouse placenta was demonstrated *in vitro*[2,21], the concomitant high level of *Maoa* expression suggests that the placenta would prevent the vasoconstrictive effect of any free maternal blood 5-HT (a small fraction of total blood 5-HT[19]) rather than transfer it to the fetus[21]. This potentially lethal bioactivity and the quasi-absence of maternal blood 5-HT transfer to the fetus has been demonstrated previously[20,22] and also in this study. A recent study concluded that maternal 5-HT is crucial for early fetal development (before E11[1]) based on abnormal phenotypes that emerge in embryos of *Tph1*$^{-/-}$ pregnant dams. Indirect effects of the *Tph1*$^{-/-}$ maternal mutation, however, cannot be excluded as *Tph1*$^{-/-}$ mice are diabetic[23], a pathological condition that affects fetal development independently of a direct maternal 5-HT effect[24]. Although the importance of a maternal source of 5-HT before and during placentation remains an open question, our results provide the first direct evidence for maternal influences on fetal brain development through a precursor that is metabolized directly by the placenta. Given our demonstration of synthetic capability

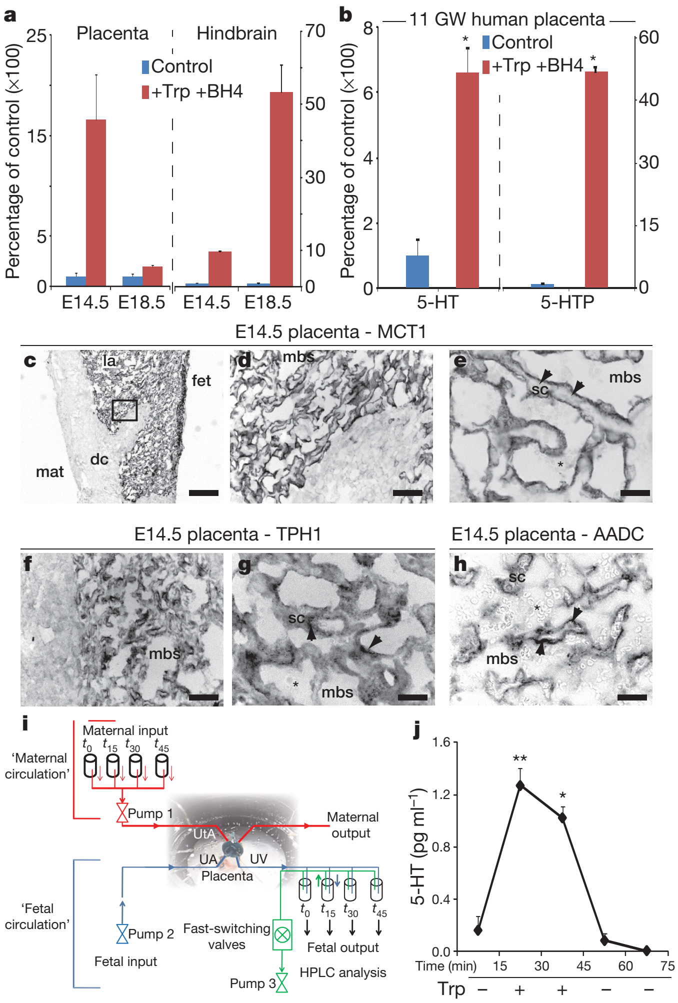


Figure 2 | Placental synthesis of 5-HT *in vitro* and *ex vivo*. **a**, Placental and hindbrain extracts were incubated with the cofactor tetrahydrobiopterin (BH4; control) or BH4 and tryptophan (+Trp+BH4) and 5-HT neosynthesis was measured after 30 min. **b**, Similar experiments conducted using human placenta tissue collected at 11 weeks of gestation (11 GW) show that human fetal villi synthesize 5-HTP and 5-HT (statistical significance versus control analysed by Student's *t*-test; *, $P < 0.005$; $n = 3$; data are presented as means ± s.e.m.). **c**, Immunostaining for the monocarboxylate transporter MCT1, a marker of syncytiotrophoblastic cells (sc) in the labyrinth (la) region on the fetal side (fet) of an E14.5 mouse placenta. On the maternal side (mat) the decidua (dc) is devoid of staining. **d–h**, Higher magnifications of the region boxed in **c**; MCT1 is expressed on the apical side (arrows) of syncytiotrophoblasts facing the maternal blood space (mbs) (**d**, **e**). The 5-HT synthetic enzymes TPH1 (**f**, **g**) and AADC (**h**) are expressed in overlapping patterns in the cytoplasm of syncytiotrophoblastic cells (arrows). Asterisks (**e**, **g**, **h**) indicate red blood cells of maternal origin. Scale bars: **c**, 20 mm; **d, f**: 500 μm; **e**, **g**, **h**: 150 μm. **i**, Schematics of the *ex vivo* dual perfusion system for the mouse placenta. UA, umbilical artery; UtA, uterine artery; UV, umbilical vein. **j**, 5-HT neosynthesis in *ex vivo* dually perfused mouse placentas at E17.5. L-tryptophan (100 μM) was injected through the uterine artery. 5-HT, neo-synthesized from maternal tryptophan and released into the umbilical vein, is evident within 15 min of precursor injection (statistical significance of 5-HT levels variation across time was analysed by one-way ANOVA; *, $P < 0.05$; **, $P < 0.01$; data from three independent experiments are presented as means ± s.e.m.).

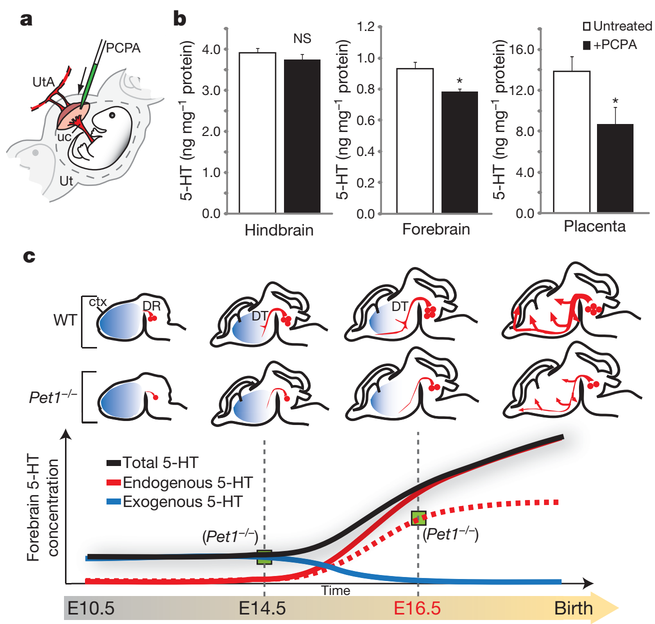


Figure 3 | HPLC measures of 5-HT concentrations in E14.5 hindbrain, forebrain and placenta of *in utero* PCPA-injected mice. **a**, Illustration of the *in vivo* injection procedure: PCPA or vehicle solution was injected into the labyrinth zone of the placenta. After 30 min, dams were killed and tissue collected and processed for HPLC. Uc, umbilical cord; Ut, uterus; UtA, uterine artery. **b**, Compared to vehicle injection (untreated, $n = 6$), placental PCPA injection (+PCPA, $n = 4$) has no significant (NS) effect on 5-HT concentration in the hindbrain, whereas it significantly lowers 5-HT concentration in the forebrain and the placenta (statistical significance versus 5-HT levels in untreated tissue was analysed by Student's *t*-test; *, $P < 0.05$; data are presented as means ± s.e.m.). **c**, Model of the progressive switch of the source of 5-HT in the fetal forebrain, from an early exogenous (placental, blue line) to a later endogenous (5-HT axons, red line) source. The green boxes represent the amount of 5-HT measured in the *Pet1*$^{-/-}$ mice, which lack most of the endogenous neuronal source. ctx, cortex; DR, dorsal raphe; DT, dorsal thalamus.

in the early human placenta, it will be clinically important to define the specific time period during human pregnancy for this placental influence on brain development. The present results also emphasize the need to examine fetal and placental tryptophan availability. Mutations in tryptophan 2,3-dioxygenase degrading enzymes (TDO1 and TDO2, which are expressed in the placenta[25]) affect neurogenesis, produce anxiety-related behaviour in mice[26] and are associated with increased risks of schizophrenia, bipolar disorder and autism[27,28]. Our results provide a mechanism through which alterations of tryptophan metabolic pathways in the placenta would affect placental 5-HT synthesis and fetal forebrain development.

Our study provides a new framework for understanding the complexity of maternal-fetal relationships that can influence brain structure and function. We focused on the 5-HT system and showed that there is a progressive switch from an early dependence on an exogenous (placental) source of 5-HT to a later endogenous brain source (Fig. 3c). The exogenous source of 5-HT is provided to the forebrain during developmental epochs that include cortical neurogenesis, migration and initial axon targeting[10]. These events can be negatively affected by disrupting 5-HT signalling: we demonstrated that a forebrain-specific disruption of 5-HT signalling *in vivo* affects axon guidance, leading to abnormal thalamocortical axon trajectories[4]. This phase in the mouse corresponds to the first and early second trimesters in the human: prenatal periods that are associated with greater risk for mental illnesses due to maternal perturbations[29,30]. Thus, translation of our findings to those corresponding periods in the human will be of significant clinical relevance.

METHODS SUMMARY

Animals and reagents. Timed-pregnant CD-1 mice were purchased from the Charles River Laboratory. Plug date was considered E0.5 and the age of individual embryos was confirmed by measuring the crown–rump length and checking for developmental landmarks such as digits and eye formation. *Pet1* (ref. 3) knockout ($-/-$), heterozygotes ($+/-$) and wild-type ($+/+$) littermate embryos were generated by crossing *Pet1*$^{+/-}$ males and females. *Sert*$^{+/-}$ embryos were obtained by crossing *Sert*$^{-/-}$ females[19] with wild-type C57BL/6 males. The *Pet1* and *Sert* knockout lines have been backcrossed on the C57Bl/6J background for $>$10 generations. The *Maoa* spontaneous knockout mouse (s*Maoa* KO) was described earlier[16]. All procedures using mice were approved by the Institutional Animal Care and Use Committee at University of Southern California and conformed to NIH guidelines. Unless otherwise noted, all reagents and antibodies were purchased from Sigma. Protocols for HPLC, tissue staining, *in vitro* 5-HT synthesis assays, MAOA enzymatic activity assays, *in vivo* synthesis inhibition assays and *ex vivo* placental perfusions are described in Methods.

Human placenta samples. Normal human placental villi samples were obtained from women undergoing elective pregnancy terminations at the Reproductive Options Clinic at the Los Angeles County (LAC) and University of Southern California (USC) Medical Center. For the purposes of this study, the tissue was carefully dissected by the pathologist and immediately flash-frozen in liquid nitrogen. No patient characteristics other than gestational age were recorded. The University of Southern California Health Science Campus Institutional Review Board approved this study.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 August 2010; accepted 24 February 2011.

1. Cote, F. *et al.* Maternal serotonin is crucial for murine embryonic development. *Proc. Natl Acad. Sci. USA* **104,** 329–334 (2007).
2. Yavarone, M. S., Shuey, D. L., Sadler, T. W. & Lauder, J. M. Serotonin uptake in the ectoplacental cone and placenta of the mouse. *Placenta* **14,** 149–161 (1993).
3. Hendricks, T. J. *et al.* Pet-1 ETS gene plays a critical role in 5-HT neuron development and is required for normal anxiety-like and aggressive behavior. *Neuron* **37,** 233–247 (2003).
4. Bonnin, A., Torii, M., Wang, L., Rakic, P. & Levitt, P. Serotonin modulates the response of embryonic thalamocortical axons to netrin-1. *Nature Neurosci.* **10,** 588–597 (2007).
5. Ansorge, M. S., Zhou, M., Lira, A., Hen, R. & Gingrich, J. A. Early-life blockade of the 5-HT transporter alters emotional behavior in adult mice. *Science* **306,** 879–881 (2004).
6. Gross, C. *et al.* Serotonin1A receptor acts during development to establish normal anxiety-like behaviour in the adult. *Nature* **416,** 396–400 (2002).
7. Taylor, S. E. *et al.* Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. *Biol. Psychiatry* **60,** 671–676 (2006).
8. Bonnin, A., Peng, W., Hewlett, W. & Levitt, P. Expression mapping of 5–HT1 serotonin receptor subtypes during fetal and early postnatal mouse forebrain development. *Neuroscience* **141,** 781–794 (2006).
9. Anderson, G. M. *et al.* Serotonin transporter promoter variants in autism: functional effects and relationship to platelet hyperserotonemia. *Mol. Psychiatry* **7,** 831–836 (2002).
10. Gaspar, P., Cases, O. & Maroteaux, L. The developmental role of serotonin: news from mouse molecular genetics. *Nature Rev. Neurosci.* **4,** 1002–1012 (2003).
11. Buznikov, G. A., Lambert, H. W. & Lauder, J. M. Serotonin and serotonin-like substances as regulators of early embryogenesis and morphogenesis. *Cell Tissue Res.* **305,** 177–186 (2001).
12. Aitken, A. R. & Tork, I. Early development of serotonin-containing neurons and pathways as seen in wholemount preparations of the fetal rat brain. *J. Comp. Neurol.* **274,** 32–47 (1988).
13. Lidov, H. G. & Molliver, M. E. Immunohistochemical study of the development of serotonergic neurons in the rat CNS. *Brain Res. Bull.* **9,** 559–604 (1982).
14. Lebrand, C. *et al.* Transient uptake and storage of serotonin in developing thalamic neurons. *Neuron* **17,** 823–835 (1996).
15. Lebrand, C. *et al.* Transient developmental expression of monoamine transporters in the rodent forebrain. *J. Comp. Neurol.* **401,** 506–524 (1998).
16. Scott, A. L., Bortolato, M., Chen, K. & Shih, J. C. Novel monoamine oxidase A knock out mice with human-like spontaneous mutation. *Neuroreport* **19,** 739–743 (2008).
17. Lynn-Bullock, C. P., Welshhans, K., Pallas, S. L. & Katz, P. S. The effect of oral 5-HTP administration on 5-HTP and 5-HT immunoreactivity in monoaminergic brain regions of rats. *J. Chem. Neuroanat.* **27,** 129–138 (2004).
18. Branchek, T. A. & Gershon, M. D. Time course of expression of neuropeptide Y, calcitonin gene-related peptide, and NADPH diaphorase activity in neurons of the developing murine bowel and the appearance of 5-hydroxytryptamine in mucosal enterochromaffin cells. *J. Comp. Neurol.* **285,** 262–273 (1989).
19. Chen, J. J. *et al.* Maintenance of serotonin in the intestinal mucosa and ganglia of mice that lack the high-affinity serotonin transporter: Abnormal intestinal motility and the expression of cation transporters. *J. Neurosci.* **21,** 6348–6361 (2001).
20. Howd, R. A., Nelson, M. F. & Lytle, L. D. L-tryptophan and rat fetal brain serotonin. *Life Sci.* **17,** 803–811 (1975).
21. Balkovetz, D. F., Tiruppathi, C., Leibach, F. H., Mahesh, V. B. & Ganapathy, V. Evidence for an imipramine-sensitive serotonin transporter in human placental brush-border membranes. *J. Biol. Chem.* **264,** 2195–2198 (1989).
22. Robson, J. M. & Senior, J. B. The 5-hydroxytryptamine content of the placenta and foetus during pregnancy in mice. *Br. J. Pharmacol. Chemother.* **22,** 380–391 (1964).
23. Kim, H. *et al.* Serotonin regulates pancreatic beta cell mass during pregnancy. *Nature Med.* **16,** 804–808 (2010).
24. Jawerbaum, A. & White, V. Animal models in diabetes and pregnancy. *Endocr. Rev.* **31,** 680–701 (2010).
25. Suzuki, S. *et al.* Expression of indoleamine 2,3-dioxygenase and tryptophan 2,3-dioxygenase in early concepti. *Biochem. J.* **355,** 425–429 (2001).
26. Kanai, M. *et al.* Tryptophan 2,3-dioxygenase is a key modulator of physiological neurogenesis and anxiety-related behavior in mice. *Mol. Brain* **2,** 8 doi:10.1186/1756-6606-2-8 (2009).
27. Miller, C. L. *et al.* Two complex genotypes relevant to the kynurenine pathway and melanotropin function show association with schizophrenia and bipolar disorder. *Schizophr. Res.* **113,** 259–267 (2009).
28. Nabi, R., Serajee, F. J., Chugani, D. C., Zhong, H. & Huq, A. H. Association of tryptophan 2,3 dioxygenase gene polymorphism with autism. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **125B,** 63–68 (2004).
29. Brown, A. S. & Derkits, E. J. Prenatal infection and schizophrenia: a review of epidemiologic and translational studies. *Am. J. Psychiatry* **167,** 261–280 (2010).
30. Oberlander, T. F., Gingrich, J. A. & Ansorge, M. S. Sustained neurobehavioral effects of exposure to SSRI antidepressants during development: molecular to clinical evidence. *Clin. Pharmacol. Ther.* **86,** 672–677 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H.H. Wu and K. Eagleson for discussions and comments on the manuscript. We thank R. Johnson (Vanderbilt HPLC core facility), L. Zhang for technical help and E. Meng for discussions. This work was supported by the NICHD (grant 5R21HD065287 to A.B.), NARSAD (A.B.) and the NIMH (grant R01MH39085 to J.C.S. and 1P50MH078280A1 to R.D.B. and P.L.).

Author Contributions A.B. conducted the experiments with assistance from N.G. in placenta studies, K.C. and J.C.S. in providing mutant mouse strains and in MAOA enzymatic assays. R.D.B. and E.S.D. provided mutant mouse strains and M.L.W. and J.K. provided human tissue. A.B. and P.L. conceived this study, interpreted the data and wrote the manuscript. All authors commented on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.B. (bonnin@usc.edu) or P.L. (plevitt@usc.edu).

METHODS

Animals and reagents. Timed-pregnant CD-1 mice were purchased from the Charles River Laboratory. Plug date was considered E0.5 and the age of individual embryos was confirmed by measuring the crown–rump length and checking for developmental landmarks such as digits and eye formation. *Pet1* (ref. 3) knockout (−/−), heterozygotes (+/−) and wild-type (+/+) littermate embryos were generated by crossing $Pet1^{+/-}$ males and females. $Sert^{+/-}$ embryos were obtained by crossing $Sert^{-/-}$ females[19] with wild-type C57BL/6 males. The *Pet1* and *Sert* knockout lines have been backcrossed on the C57Bl/6J background for >10 generations. The *Maoa* spontaneous knockout mouse (s*Maoa* KO) was described earlier[16]. All procedures using mice were approved by the Institutional Animal Care and Use Committee at University of Southern California and conformed to NIH guidelines. Unless otherwise noted, all reagents and antibodies were purchased from Sigma.

Human placenta samples. Normal human placental villi samples were obtained from women undergoing elective pregnancy terminations at the Reproductive Options Clinic at the Los Angeles County (LAC) and University of Southern California (USC) Medical Center. For the purposes of this study, the tissue was carefully dissected by the pathologist and immediately flash-frozen in liquid nitrogen. No patient characteristics other than gestational age were recorded. The University of Southern California Health Science Campus Institutional Review Board approved this study.

HPLC measures. Hindbrains and forebrains were homogenized, using a tissue dismembrator, in 100–750 µl of 0.1 M trichloroacetic acid TCA, which contains 10^{-2} M sodium acetate, 10^{-4} M EDTA, 5 ng ml^{-1} isoproterenol (as internal standard) and 10.5% methanol (pH 3.8). Samples were spun at 10,000*g* for 20 min. The supernatant was removed and stored at −80 °C. The pellet was used for protein analysis; total protein concentrations were determined using the BCA Protein Assay Kit (Pierce Chemical Company). The supernatant was then thawed and spun for 20 min. The concentration of biogenic amines in each sample was determined by a specific HPLC assay using an Antec Decade II (oxidation, 0.5 V) electrochemical detector operated at 33 °C. 20 µl samples of supernatant were injected using a Water 717+ autosampler onto a PhenomenexNucleosil (5 u, 100 A) C18 HPLC column (150 × 4.60 mm). Biogenic amines were eluted with a mobile phase consisting of 89.5% 100 mM TCA, 10 mM sodium acetate, 0.1 mM EDTA and 10.5% methanol (pH 3.8). Solvent was delivered at 0.6 ml min^{-1} using a Waters 515 HPLC pump. Using this HPLC solvent, the biogenic amines eluted in the following order: noradrenaline, adrenaline, DOPAC, dopamine, 5-HIAA, HVA, 5-HT and 3-MT. HPLC control and data acquisition were managed by Millennium 32 software.

Statistical analysis. Measures in $Pet1^{-/-}$ and wild-type littermate embryos ($n \geq 6$ per genotype) are presented as mean ± s.e.m. Statistical comparison of measures between genotypes across development was performed using a one-way ANOVA (Fig. 1). Differences were considered statistically significant at $P < 0.05$. Single statistical comparison between two groups was performed using a two-tailed Student's *t*-test (for example, comparison of measures in $Pet1^{-/-}$ and wild-type littermates or s$Maoa^{-/-}$ and wild-type littermates at a single age or *in utero* PCPA- versus vehicle-injected mice; see Fig. 3 and Supplementary Tables 1 and 2).

***In vitro* 5-HT synthesis assays.** Pregnant CD-1 mice were anaesthetized by an intraperitoneal injection of 0.2 ml pentobarbital and perfused through the left ventricle with 100 ml of phosphate buffered saline (PBS). The uterus was immediately dissected and the resulting embryos were placed on ice in PBS. The placenta, forebrain and hindbrain were removed and kept on ice in 1.5 ml eppendorf tubes. The collected tissue was homogenized in 3 volumes of 0.05 M Tris buffer, pH 7.5, containing 1 mM dithiothreitol and 1 mM EGTA. Homogenates were centrifuged at 29,000*g* for 15 min at 4 °C and the supernatants were placed on ice; tryptophan hydroxylase assays were performed as previously described[31]. Briefly, 70 µg of protein from the supernatant was added to tubes containing 100 µl of: 0.05 M Tris buffer, pH 7.5, 1 mM EGTA, 15 µg catalase, 250 µM L-tryptophan, and 50 µM tetrahydrobiopterin (BH4; a cofactor required for TPH1 and TPH2 activity[32]). Reactions were performed for 10 and 30 min at 37 °C and stopped by flash freezing in liquid nitrogen. L-tryptophan metabolism was determined by measuring 5-hydroxytryptophan (5-HTP), 5-hydroxytyptamine (5-HT), and 5-hydroxyindoleacetic acid (5-HIAA) using HPLC with electrochemical detection (see above). Protein concentrations were determined using a DC protein assay (Bio-Rad).

***In vitro* MAO-A activity measures.** MAO-A activity was determined by incubating forebrain and hindbrain homogenates with [^{14}C]-5-HT (1 mM) at 37 °C for 20 min. At the end of the incubation, the products were extracted and the radioactivity determined as described previously[33].

Immunohistochemistry. Brain sections were processed for immunohistochemistry using the following primary antibodies: anti-5-HT (rabbit, 1:1,000; goat, 1:1,000, Sigma), anti-NetrinG1a (goat, 1:250, RnD), anti-SERT (rabbit, 1:200, Sigma), anti-5-HTP (rabbit, 1:1,500; ImmunoStar) and anti-AADC (rabbit, 1:500, Abcam). The secondary antibodies (Jackson ImmunoResearch) were: Cy2- or Cy3-conjugated donkey anti-rabbit IgG (1:800), rhodamin-X conjugated donkey anti-goat (1:200) and biotin-conjugated anti-rabbit (1:1,000). 20 µm brain sections were permeabilized, and pre-incubated with blocking solution (0.1% Triton X-100, 2% fetal bovine serum) in PBS for 2 h and then incubated overnight with a primary antibody diluted in the same blocking solution at 4 °C. Sections were washed with PBS + 0.1% Triton X-100 and incubated with appropriate secondary antibodies diluted in same blocking solutions for a minimum of 2 h at room temperature (22-25 °C). Sections were then washed with PBS + 0.1% Triton X-100 and mounted using a fluorescence anti-fade medium (Prolong Gold, Vector Laboratories). For non-fluorescent immunohistochemistry (for example, AADC staining), similar procedures were followed, except that the sections were incubated in methanol with 3% H_2O_2 for 15 min to quench endogenous peroxidase activity. The reaction product was visualized using a Vector ABC Elite Kit (Vector Laboratories). Images were acquired on a Zeiss AxoplanII microscope coupled to an AxioCamHRc camera using Axiovision 4.1 software (Zeiss). For placental immunochemistry, 5 µm frozen sections were thawed for 15 min at room temperature then washed in a mixture of 1× PBS containing 0.1% Triton X-100 (PBS-T) for 15 min at room temperature to allow for membrane permeabilization. Endogenous HRP was quenched by incubating the slides in a solution of methanol containing 3% hydrogen peroxide for 15 min at room temperature. The slides were then incubated for 2 h in a solution of PBS-T containing 2% Fetal Bovine Serum (Blocking Solution). The slides were then incubated overnight with anti-MCT1 (MCT1 is a marker of syncytiotrophoblasts[34]; chicken; 0.5 µg ml^{-1}) or anti-AADC or anti-TPH1 (rabbit; 1:500) overnight at room temperature. Slides were then washed 3 × 10 min in PBS-T and incubated with appropriate secondary antibodies (biotin-conjugated anti-chicken; 1:1,000 or biotin-conjugated anti-rabbit; 1:1,000; Jackson ImmunoResearch) for 2 h at room temperature. After 6 × 5 min washes, slides were processed for HRP visualization as described above.

***In utero* inhibition of placental TPH1.** Pregnant CD-1 mice were anaesthetized by an intraperitoneal injection of 0.5–0.7 ml of 10% Nembutal and uterine horns exposed as previously described[4]. Glass micropipettes were filled with a solution of 10 mM PCPA (4-Chloro-DL-phenylalanine; an inhibitor of TPH1 enzymatic activity[35]) in PBS and 0.1% fast green dye (FCF; Sigma). FCF was used to monitor visually the diffusion of solution into the fetal villi of the placenta. 2–5 µl were injected through the uterine wall into the labyrinth region in half of the placentas ($n = 4$–6) for each pregnant mouse ($n = 3$); remaining placentas were injected with vehicle only. After 30 min, the dam was killed and placentas and corresponding forebrains and hindbrains collected and quickly frozen in liquid N_2. Samples were then processed for HPLC analysis as described above.

***Ex vivo* dual perfusion of the mouse placenta.** A single placenta with uterine and umbilical arteries and veins attached was harvested from a pregnant CD-1 mouse at E17.5 in warm PBS. The placenta was rapidly transferred to a thermostated tissue culture chamber (60 mm diameter, Warner Instruments) to which warm (37 °C) and oxygenated (95% O_2, 5% CO_2) PBS solution was constantly provided. Using a dissection microscope, micro-catheters were carefully attached to the uterine artery on the maternal side and umbilical artery and vein on the fetal side of the placenta. The uterine artery at E17.5 (internal diameter is approximately 150 µm) was cannulated using a 150 µm outside diameter polyethylene micro-catheter. Micro-catheters were kept solidly attached to the arteries or veins using micro-clamps (F.S.T.) and suture. The other end of the uterine artery micro-catheter was connected to an 8-channel micro-manifold (Automate Scientific); each channel's polyethylene tubing was then connected to an individual fast-switching valve (1.5 ms response time; Automate Scientific); the artificial maternal solution (warm oxygenated PBS) was delivered to each individual channel through a pressurized cylinder connected to each valve. An electronic valve controller enabled timed sequential switching of maternal solutions. The input pressure was calibrated to allow a maximal flow rate of 10–20 µl min^{-1} through the uterine artery micro-catheter. On the fetal side, the umbilical artery and vein were similarly cannulated; for E17.5 placentas polyethylene tubing with 100 µm internal diameter was used. The other end of the umbilical artery micro-catheter was directly connected to a 5 ml syringe containing the artificial fetal solution (PBS with 0.01% fast green dye FCF, Sigma) and mounted on a precision, computer-controlled syringe pump (Braintree Scientific); the infusion flow rate was set at 6 µl min^{-1}. For fetal eluate collection, the umbilical vein was cannulated to 100 µm internal diameter polyethylene tubing; this tubing was then connected to a specially modified SmartSquirt system (Automate Scientific) which consisted of a micro-manifold connected to eight independent 3-ml reservoirs in which a controlled negative pressure (−15 PSI) was individually applied. The integrated fast-switching valves system allowed the application of negative pressure to each reservoir

independently and sequentially, thus allowing individual fraction collection over time. The addition of fast green dye to the fetal perfusate allowed monitoring for any leak through and around the placenta during the perfusion.

31. Johansen, P. A., Jennings, I., Cotton, R. G. & Kuhn, D. M. Tryptophan hydroxylase is phosphorylated by protein kinase A. *J. Neurochem.* **65,** 882–888 (1995).
32. Kuhn, D. M., Ruskin, B. & Lovenberg, W. Tryptophan hydroxylase. The role of oxygen, iron, and sulfhydryl groups as determinants of stability and catalytic activity. *J. Biol. Chem.* **255,** 4137–4143 (1980).
33. Chen, K. *et al.* Forebrain-specific expression of monoamine oxidase A reduces neurotransmitter levels, restores the brain structure, and rescues aggressive behavior in monoamine oxidase A-deficient mice. *J. Biol. Chem.* **282,** 115–123 (2007).
34. Nagai, A., Takebe, K., Nio-Kobayashi, J., Takahashi-Iwanaga, H. & Iwanaga, T. Cellular expression of the monocarboxylate transporter (MCT) family in the placenta of mice. *Placenta* **31,** 126–133 (2010).
35. Koe, B. K. & Weissman, A. p-Chlorophenylalanine: a specific depletor of brain serotonin. *J. Pharmacol. Exp. Ther.* **154,** 499–516 (1966).

doi:10.1038/nature09865

Neuronal activity is required for the development of specific cortical interneuron subtypes

Natalia V. De Marco García[1]*, Theofanis Karayannis[1]* & Gord Fishell[1]

Electrical activity has been shown to regulate development in a variety of species and in various structures[1], including the retina[2–4], spinal cord[5,6] and cortex[5]. Within the mammalian cortex specifically, the development of dendrites and commissural axons in pyramidal cells is activity-dependent[7,8]. However, little is known about the developmental role of activity in the other major cortical population of neurons, the GABA-producing interneurons. These neurons are morphologically and functionally heterogeneous and efforts over the past decade have focused on determining the mechanisms that contribute to this diversity[9–11]. It was recently discovered that 30% of all cortical interneurons arise from a relatively novel source within the ventral telencephalon, the caudal ganglionic eminence (CGE)[11,12]. Owing to their late birth date, these interneurons populate the cortex only after the majority of other interneurons and pyramidal cells are already in place and have started to functionally integrate. Here we demonstrate in mice that for CGE-derived reelin (Re)-positive and calretinin (Cr)-positive (but not vasoactive intestinal peptide (VIP)-positive) interneurons[12,13], activity is essential before postnatal day 3 for correct migration, and that after postnatal day 3, glutamate-mediated activity controls the development of their axons and dendrites. Furthermore, we show that the engulfment and cell motility 1 gene (*Elmo1*)[14], a target of the transcription factor distal-less homeobox 1 (*Dlx1*)[15], is selectively expressed in Re^+ and Cr^+ interneurons and is both necessary and sufficient for activity-dependent interneuron migration. Our findings reveal a selective requirement for activity in shaping the cortical integration of specific neuronal subtypes.

Experimental evidence indicates that interneurons are electrically active shortly after their birth and participate in the early network activity that may contribute to circuit maturation in the neonatal cortex[16–18]. However, the role of activity in developing interneuron subtypes has not been addressed. Here we demonstrate that altering the level of neuronal excitability *in vivo* within genetically targeted CGE-derived interneurons has profound consequences on multiple aspects of the development of select subtypes within this population, as well as their associated gene expression (Supplementary Fig. 1).

To suppress neuronal excitability within CGE-derived interneurons, we electroporated *in utero* the inward rectifying potassium channel Kir2.1 under the control of the *Dlx5/6* enhancer element[19] at embryonic day (E)15.5, which results in selective expression within CGE-derived interneuron populations (Supplementary Fig. 2). Kir2.1-overexpression has been shown to affect activity by lowering the resting membrane potential (V_{rest}), therefore altering neuronal excitability[20]. We detected expression of this channel by *in situ* hybridization (Supplementary Fig. 3a, b). To functionally assess the presence of membrane-targeted channels, we performed whole-cell patch-clamp recordings from *Dlx5/6-Kir2.1*, *Dlx5/6-eGFP* co-electroporated interneurons in voltage clamp at P8–P9. Current/voltage (I/V) curve analysis indicated the presence of an inward rectifying potassium conductance that was active at V_{rest} and was blocked by 300 μM barium, a concentration that preferentially blocks Kir2.1 channels (Supplementary Fig. 3d–g). Consistent with these observations, the V_{rest} of Kir2.1-electroporated interneurons was significantly more hyperpolarized than that of interneurons electroporated with eGFP alone (Supplementary Fig. 3c).

By postnatal day (P)8, subsets of the interneurons expressing the Kir2.1 channel showed pronounced defects in their morphologies (Fig. 1 and Supplementary Figs 4a, b, 5). To quantify alterations in dendrites and axons, we reconstructed interneuron morphologies from cortical slices at P8–P9, the earliest stages at which interneuron subtypes can be consistently delineated by expression of immunochemical markers. Our analysis revealed that the total length of axonal arborizations was significantly reduced in multipolar and bipolar Cr^+ interneurons, as well as neurogliaform and dense plexus Re^+ subtypes, whereas those of multipolar VIP^+ interneurons remained unaltered (Fig. 1 and Supplementary Fig. 4a, b). Quantification of axonal nodes and ends (see Methods) also revealed scantily branched axons in Cr^+ and Re^+ subtypes but not in VIP^+ interneurons (Fig. 1b and Supplementary Fig. 5a–c). Although total dendritic length was not significantly decreased in Re^+ interneurons (Fig. 1c), this subtype exhibited less complex dendritic trees (Fig. 1c and Supplementary Fig. 5f). In contrast, VIP^+ and Cr^+ interneurons showed normal dendritic morphologies (Fig. 1c and Supplementary Fig. 5d, e). To assess whether the morphological defects observed in Cr^+ and Re^+ interneurons were due to a developmental delay, we analysed the electrophysiological properties and morphology of these interneurons at P15–19. Despite possessing mature intrinsic properties, Cr^+ and Re^+ interneurons at P15–19 showed morphological defects similar to those found at P8 (Supplementary Fig. 4c–h). These findings indicate that the observed defects are unlikely to be simply a result of a developmental delay in maturation.

Although neuronal activity has been shown to be dispensable for the migration of pyramidal cells[7], we noticed a pronounced overall shift in the laminar positioning of CGE subtypes expressing Kir2.1. CGE-derived interneurons migrate tangentially from the ventral telencephalon to the cortex where they then undergo radial migration to reach stereotypic positions in cortical laminae by P7. To assess the role of neuronal activity during interneuron migration, we used a *tetO-Kir2.1.ires.LacZ* transgenic mouse line in which Kir2.1 and LacZ are expressed on binding of the tet transactivator (Tta) to the *tetO* element[20]. We electroporated a *Dlx5/6-Tta* plasmid together with *Dlx5/6-eGFP*, again at E15.5 (Supplementary Fig. 6a), to induce Kir2.1 expression selectively in CGE-derived interneurons. These experiments revealed that the tangential migration of interneurons expressing Kir2.1 was indistinguishable from control populations at early developmental stages (Fig. 2a). However, after P5, and in agreement with the constitutive *Dlx5/6-Kir2.1* electroporation experiment, interneurons that expressed Kir2.1 were found to occupy deeper cortical layers than control populations (Fig. 2a). To analyse the selectivity of this defect, we quantified the distribution of Cr^+, Re^+ and VIP^+ interneurons across all cortical layers. We detected a significantly higher percentage of Kir2.1 Cr^+ interneurons in layer IV and a concomitant

[1]Smilow Neuroscience Program, Departments of Cell Biology and Neural Science, New York University Langone Medical Center, New York, New York 10016, USA.
*These authors contributed equally to this work.

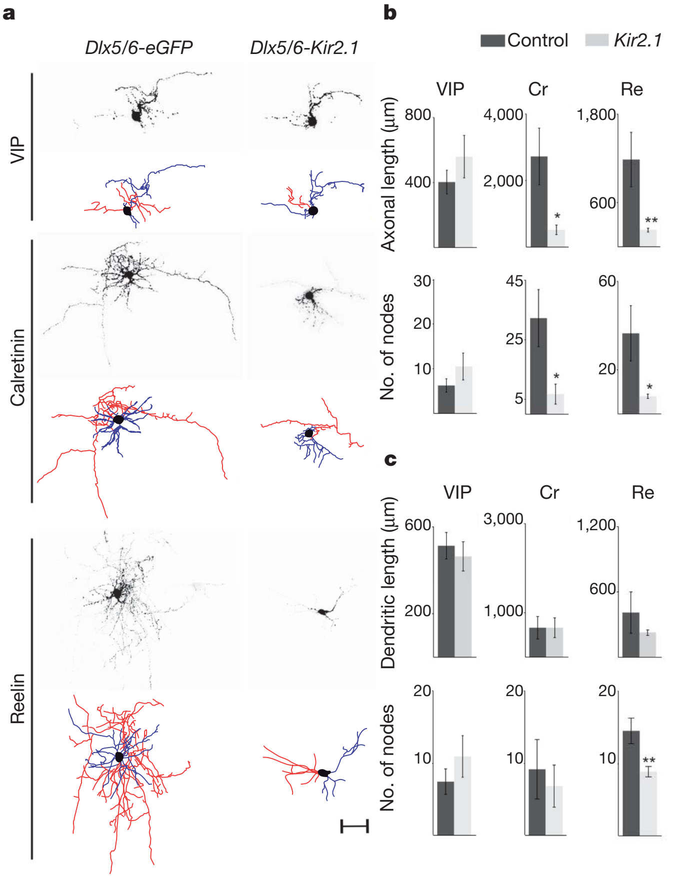


Figure 1 | Defective morphology of Cr^{+} and Re^{+} interneuron subtypes resulting from Kir2.1 expression. **a**, Representative examples of P8 VIP^{+}, Cr^{+} and Re^{+} interneurons in mice electroporated at E15.5 with *Dlx5/6-eGFP* (control) or *Dlx5/6-eGFP*, *Dlx5/6-Kir2.1* plasmids at E15.5. Photomicrographs of eGFP expression and corresponding neurolucida reconstructions depicting axons (red), dendrites (blue) and somata (black). Scale bar, 50 µm. **b**, Morphometric analysis of control and Kir2.1-expressing VIP^{+}, Cr^{+} and Re^{+} subtypes including the total length of axonal arborizations (top) and number of axonal nodes (bottom). **c**, Total length of dendritic trees (top) and number of dendritic nodes (bottom) in the same subtypes. Mean values (± s.e.m.) were obtained from >4 reconstructed interneurons each in *Dlx5/6-eGFP* and *Dlx5/6-eGFP*, *Dlx5/6-Kir2.1* electroporated mice. Paired *t*-test: $*P < 0.05$, $**P < 0.01$.

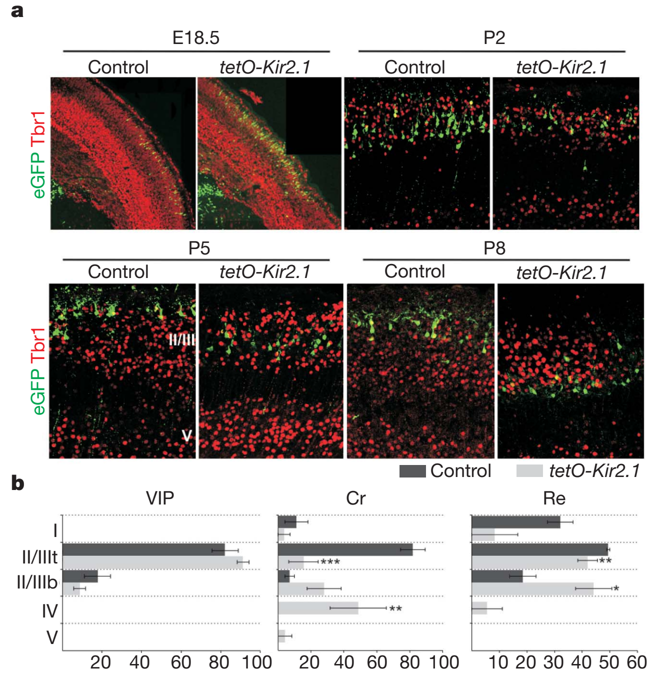


Figure 2 | Neuronal activity is essential for the proper laminar migration of selective interneuron subtypes. **a**, Laminar positioning of electroporated interneurons in wild-type mice (control) and *tetO-Kir2.1.ires.LacZ* littermates both co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids at E15.5. Tbr1 expression delineates layers II/III and V at P5–P8. Representative examples taken from the analysis of four control and six *tetO-Kir2.1.ires.LacZ* electroporated mice for each developmental stage. **b**, Quantification of the distribution of VIP^{+}, Cr^{+} and Re^{+} interneuron subtypes across cortical layers at P8. Owing to the lack of selective molecular markers to distinguish between cortical layer II and III at P8–P9, we divided these layers collectively into II/IIIt and II/IIIb (where II/IIIt and II/IIIb refer to II/III top and II/III bottom, respectively). Mean percentage values (± s.e.m.) were obtained from four wild-type and six *tetO-Kir2.1.ires.LacZ* electroporated mice. Paired *t*-test: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

reduction in the percentage of this population in layers II/IIIt (where II/IIIt refers to II/III, top, as layers II and III have not sorted out by this developmental timepoint) compared to controls (Fig. 2b). Similarly, in electroporated *tetO-Kir2.1.ires.LacZ* mice we observed a significantly lower percentage of Re^{+} interneurons in layer II/IIIt and a subsequent increase in layer II/IIIb (where II/IIIb refers to II/III, bottom) compared to controls (Fig. 2b). In contrast, the distribution of VIP^{+} interneurons in electroporated *tetO-Kir2.1.ires.LacZ* mice was similar to that observed in controls (Fig. 2b). Our results indicate that neuronal activity is a determinant in the allocation of Cr^{+} and Re^{+} subtypes to defined cortical layers.

One interpretation of our results is that the morphological defects observed in Cr^{+} and Re^{+} Kir2.1-expressing interneurons are an indirect consequence of the laminar mispositioning in the cortex. Alternatively, neuronal activity may regulate laminar migration and morphological maturation independently. To distinguish between these two possibilities, we took advantage of the ability of doxycycline to suppress Kir2.1 expression from the *tetO-Kir2.1.ires.LacZ* transgenic line[20] and administered it at different developmental time points (Supplementary Fig. 6a). We were able to monitor the expression of the Kir2.1 transgene by assessing β-galactosidase activity (Fig. 3b). To determine whether Kir2.1 expression had any effects on early interneuron differentiation, we treated *Dlx5/6-Tta* and *Dlx5/6-eGFP* E15.5-electroporated pregnant mice with doxycycline at E16.5. As it takes approximately three days for doxycycline administration to fully and irreversibly suppress the expression of Kir2.1 and LacZ (Supplementary Fig. 7), in these experiments Kir2.1 expression is shut off from P0 onwards. We found that Kir2.1 expression before P0 had no effect on the laminar position, immunochemical profile, morphology or intrinsic physiological properties of CGE-derived interneurons analysed at P8–P9 (Fig. 3a, data not shown). Thus, interneuron specification and maturation proceed normally if Kir2.1 is shut off by P0.

In contrast, migration defects persisted when Kir2.1 expression was shut off at P3 (Fig. 3a). Remarkably, despite their abnormal laminar position under these conditions, the morphology of Cr^{+} and Re^{+} subtypes was unperturbed (Fig. 3c, d). The total length and complexity of Cr^{+} and Re^{+} interneuron axonal arborizations was not significantly different in doxycycline-treated *tetO-Kir2.1.ires.LacZ* mice compared to wild-type controls (Fig. 3d and Supplemental Fig. 6b, d). Similarly, the complexity of the dendritic trees in Kir2.1-expressing Re^{+} interneurons after doxycycline treatment was similar to that observed in controls (Fig. 3d and Supplemental Fig. 6c). In contrast, both morphological and migratory defects persisted in *tetO-Kir2.1.ires.LacZ* mice in which Kir2.1 expression was turned off from P5 onwards (Supplementary Fig. 8, data not shown). Together these findings revealed that neuronal activity is independently required between P0 and P3 to regulate laminar position and after P3 to control the morphological development of specific interneuron subtypes.

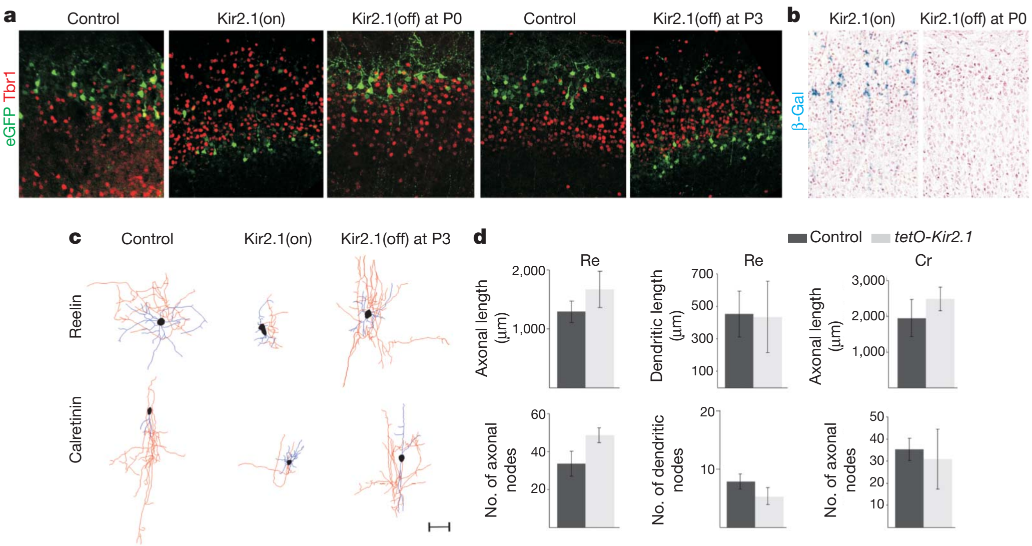


Figure 3 | Specific interneuron subtypes require activity for migration and morphological maturation at two distinct stages of development. a, Laminar positioning of P8 electroporated interneurons in wild-type mice (control) and *tetO-Kir2.1.ires.LacZ* mice both co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids at E15.5. Mice received either no treatment (Kir2.1(on)); or were treated with doxycycline at E16.5 (Kir2.1(off) at P0 onwards); or with doxycycline at P0 (Kir2.1(off) at P3 onwards). **b**, β-Galactosidase (β-Gal) activity in P8 *tetO-Kir2.1.ires.LacZ* mice co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids either untreated or treated with doxycycline at E16.5 (Kir2.1(off) at P0 onwards). **c**, Neurolucida reconstructions of Cr^+ and Re^+ interneurons in wild-type (control) and *tetO-Kir2.1.ires.LacZ* mice both co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids. Mice received either no doxycycline treatment (Kir2.1(on)) or doxycycline at P0 (Kir2.1(off) at P3 onwards). Axons are shown in red, dendrites in blue and somata in black. Scale bar, 50 μm. **d**, Quantification of dendritic and axonal morphology in control and experimental Cr^+ and Re^+ interneurons in *tetO-Kir2.1.ires.LacZ* mice after doxycycline administration at P0. Mean percentage values (± s.e.m.) were obtained from >3 reconstructed interneurons each in doxycycline-treated wild-type and *tetO-Kir2.1.ires.LacZ* mice for each subtype analysed at P8.

It is unclear what kinds of activity might be responsible for controlling these distinct aspects of subtype-specific integration at different developmental stages. Experimental evidence indicates that a large proportion of developing neurons in the central nervous system show correlated spontaneous activity[21–23]. This activity results in prominent cortical activity patterns apparent during the first postnatal week such as glutamate-dependent cortical early network oscillations[16]. Interestingly, cortical interneurons have the ability to participate in such activity as they express glutamate receptors at early stages of development[24]. To explore the possibility that interneuron maturation is regulated by glutamate-driven ionotropic receptor activity, we used kynurenic acid, an NMDA and AMPA/kainate receptor blocker[25]. We applied either kynurenic acid diluted in PBS or PBS alone (control) subdurally to the brains of *Dlx5/6-eGFP* electroporated mice at P0, P1, P2 and P3 and analysed interneuron migration and morphology at P8–P9 (Supplementary Fig. 9a and Fig. 4). Migration of all subtypes was normal after kynurenic acid injections at all ages tested (see Supplementary Information). In contrast, we observed morphological defects in Cr^+ and Re^+ subtypes in mice injected with kynurenic acid at P3 (but not after administration at earlier ages, that is, P0, P1, P2). These subtype-specific defects were reminiscent of those found in the Kir2.1 experiments (Fig. 4). Specifically, the total axonal length and complexity of Cr^+ and Re^+ interneurons was significantly reduced after kynurenic acid treatment (Fig. 4b and Supplementary Fig. 9d, f). Dendritic trees of Re^+ interneurons in kynurenic-acid-treated mice also showed a trend towards a reduction in overall length and a simplified morphology compared to controls (Fig. 4c and Supplementary Fig. 9g). In contrast, VIP^+ interneurons were not affected by kynurenic acid treatment (Fig. 4 and Supplementary Fig. 9b, c). These results indicate that ionotropic glutamate receptor-mediated activity is required after P3 to regulate the subtype-specific development of neuronal morphology but does not control their selection of cortical laminae.

To explore the molecular mechanism underlying the activity-dependent maturation of CGE-derived interneuron subtypes, we examined transcriptional programs that operate in these interneurons at early developmental stages[26,27]. Previous experimental evidence indicates that *Dlx1* is essential for both proper cortical migration and morphological development of GABAergic interneurons[15,26,28]. To determine whether *Dlx1* expression is modulated by activity, we analysed the expression of the DLX protein in control and Kir2.1-electroporated interneurons. We found that Kir2.1-expressing interneurons show lower levels of DLX expression compared to controls at P5 (Fig. 5a, c). Reduced levels of DLX expression are likely to represent attenuated *Dlx1* and/or *Dlx2* expression (see Methods). To confirm that the *Dlx1* transcriptional program is downregulated in Kir2.1-expressing interneurons, we assessed the expression of the neuronal PAS domain protein 1 (NPAS1), a previously described *Dlx1* target[28]. Consistent with a downregulation of *Dlx1*, we found that levels of NPAS1 in Re^+ subtypes were reduced upon Kir2.1 expression (Supplementary Fig. 10).

Another gene that was also shown to be a target of *Dlx* genes is *Elmo1*[15], which encodes an evolutionarily conserved Rac-activator protein[14]. We assessed ELMO1 expression in developing GABAergic interneurons (Fig. 5b), because it has been implicated in cytoskeletal reorganization and migration in the immune system[14,29]. It is also significantly downregulated in *Dlx1/2* knockout mice[15], which show severe interneuron migration defects. We found that ELMO1 is expressed by Re^+ and Cr^+ but not VIP^+ subtypes and is downregulated upon Kir2.1 expression (Fig. 5b, c). To investigate whether loss of ELMO1 function can lead to defects in interneuron migration and morphological maturation, we co-electroporated E15.5 CGE-derived interneurons with a dominant-negative form of the ELMO1 protein that impairs Rac activation, *Dlx5/6-Elmo1_TN558.Flag*[30], and *Dlx5/6-eGFP*. At P9, we detected immunoreactivity against the Flag epitope, indicating that there is robust expression of the dominant-negative protein in electroporated interneurons (Fig. 5d, inset). Interestingly, whereas electroporated Re^+ and Cr^+ interneurons show normal morphological development, these interneurons were found to be distributed within deeper layers compared to *Dlx5/6-eGFP* controls (Fig. 5d, f; data not shown). In agreement with the lack of ELMO1 expression in VIP^+ interneurons, neither their migration nor their morphology was affected by overexpression of the dominant-negative protein (data not

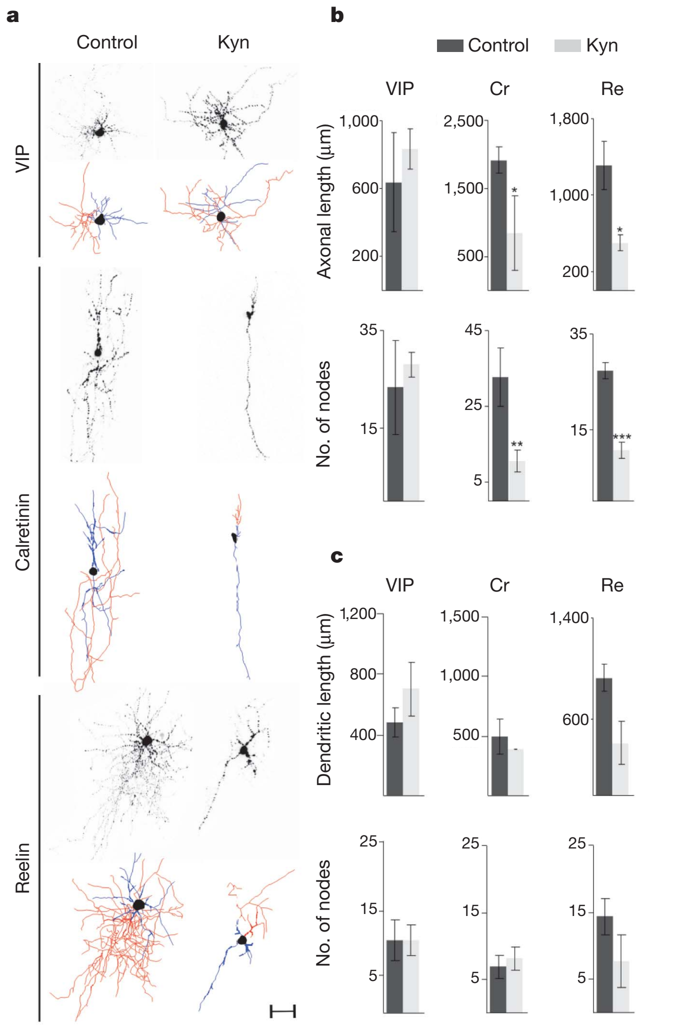


Figure 4 | Ionotropic glutamate receptor blockade mimics the effects of Kir2.1 expression on Cr^{+} and Re^{+} interneuron morphology. **a**, Representative examples of P8 VIP^{+}, Cr^{+} and Re^{+} interneurons in *Dlx5/6-eGFP* electroporated mice at E15.5 injected with PBS (control) or kynurenic acid (Kyn) at P3 and corresponding neurolucida reconstructions depicting axons (red), dendrites (blue) and somata (black). Scale bar, 50 µm. **b**, **c**, Morphometric analysis of control and kynurenic-acid-treated neurons including the total length of axonal arborizations (**b**, top) and the number of axonal nodes (**b**, bottom), and the total length of dendritic trees (**c**, top) and the number of dendritic nodes (**c**, bottom) in VIP^{+}, Cr^{+} and Re^{+} subtypes. Mean percentage values (± s.e.m.) were obtained from three electroporated interneurons each in control and kynurenic-acid-treated mice for each subtype. Paired *t*-test: $*P < 0.05$, $**P = 0.05$, $***P < 0.01$.

shown). Our observations indicate that ELMO1 is necessary for the proper radial migration of Re^{+} and Cr^{+} subtypes.

To address whether the reduction in ELMO1 expression is responsible for the abnormalities in laminar migration observed in Kir2.1-expressing Re^{+} interneurons, we co-electroporated E15.5 interneurons with a *Dlx5/6-Elmo1* construct together with *Dlx5/6-Kir2.1* and *Dlx5/6-eGFP* plasmids. We reasoned that if so the recovery of ELMO1 expression in Re^{+} and Cr^{+} Kir2.1-electroporated interneurons would rescue their migratory defects. Remarkably, the migration but not the morphology of these subtypes appeared normal in Kir2.1-electroporated interneurons that co-expressed ELMO1 at P9 (Fig. 5e, f, data not shown). As expected, neither migratory nor morphological defects were detected in VIP^{+} interneurons. In contrast, expression of *Dlx5/6-Elmo1* plasmid in the absence of *Dlx5/6-Kir2.1* did not affect migration or the morphological maturation of Re^{+}, Cr^{+} and VIP^{+} subtypes (data not shown). These results indicate that ELMO1 is necessary and sufficient for the proper

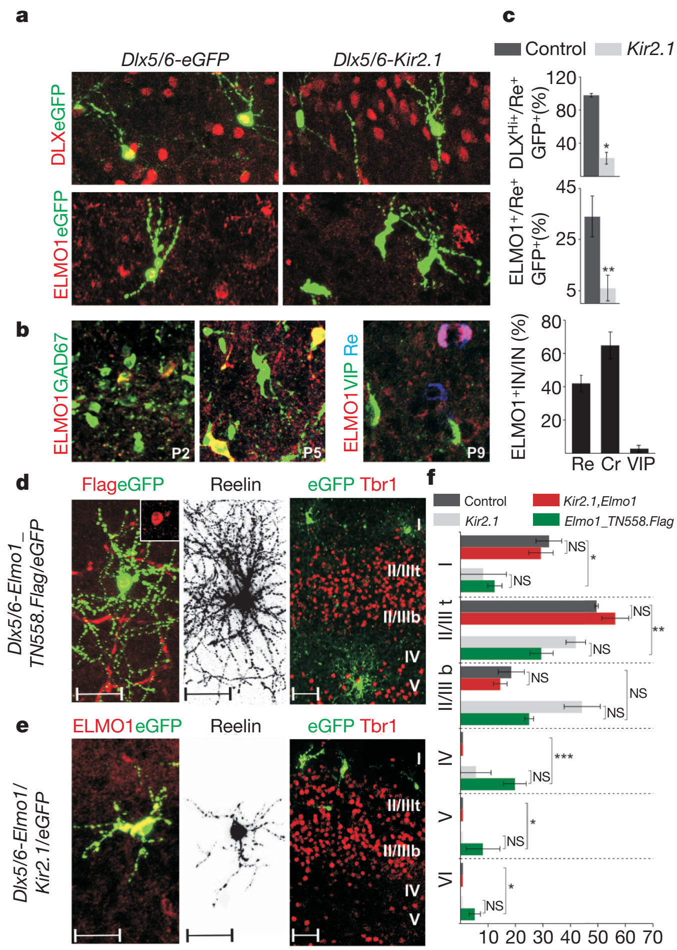


Figure 5 | Activity-dependent expression of ELMO1 regulates CGE-derived interneuron migration. **a**, Expression of *Dlx* genes and ELMO1 at P5 in *Dlx5/6-eGFP* and *Dlx5/6-Kir2.1* electroporated interneurons at E15.5. **b**, Expression of ELMO1 in *GAD67-GFP* transgenic mice at P2 and P5. Selective expression of ELMO1 in CGE-derived interneuron subtypes at P9. Quantification of ELMO1 expression in Re^{+}, Cr^{+} and VIP^{+} interneurons (IN) at P9 (right). Mean percentage values (± s.e.m.) were obtained from >70 interneurons for each subtype. **c**, Quantification of DLX^{Hi} (where Hi refers to high level of DLX protein expression) and ELMO1 expression in *Dlx5/6-eGFP* (control) and *Dlx5/6-Kir2.1* Re^{+} E15.5 electroporated interneurons at P5. Mean percentage values (± s.e.m.) were obtained from >20 interneurons each in control and Kir2.1 electroporated mice for each quantification. **d**, Electroporation of *Dlx5/6-Elmo1_TN558.Flag* plasmid at E15.5. Flag immunoreactivity is detected in electroporated interneurons at P9 (inset). Neuronal morphology of a Re^{+} interneuron and laminar distribution of electroporated interneurons at P9. Representative examples from four electroporated mice. **e**, Co-electroporation of *Dlx5/6-Elmo1* and *Dlx5/6-Kir2.1* plasmids at E15.5. ELMO1 expression in electroporated interneurons at P9. Morphological defects of an electroporated Re^{+} interneuron and laminar distribution of electroporated interneurons. Representative examples from six electroporated mice. Scale bars for **d** and **e**, 50 µm. **f**, Quantification of the distribution of Re^{+} interneurons across cortical layers at P9 on expression of different plasmids. Mean percentage values (± s.e.m.) were obtained from >80 interneurons for each group. Values for control and *Dlx5/6-Kir2.1* alone groups are repeated from Fig. 2 to facilitate comparison between groups. The large bracket indicates comparison between the control and *Dlx5/6-Elmo1_TN558.Flag* electroporated internerneurons. Paired *t*-test: $*P < 0.05$, $**P < 0.01$, $***P < 0.0001$.

activity-dependent migration of select interneuron subtypes. Taken together, our results indicate that the molecular machinery directing the maturation of Re^{+} and Cr^{+} interneurons, including *Dlx1*, *Npas1* and *Elmo1*, has evolved to be controlled by activity during development.

A role for *Dlx* genes in both interneuron migration and morphological development has been previously reported[15]; however, a link between *Dlx*

expression and neuronal activity has not been established. Our studies indicate that *Dlx1* expression and associated downstream targets are selectively regulated by activity in at least some interneuron subtypes. Specifically, *Dlx* genes induce the expression of *Elmo1*[15], which is required for proper laminar migration of Re^{+} and Cr^{+} subtypes. Although we currently provide only correlative evidence for the link between activity-regulated expression of *Dlx1* and morphological development, the alteration in interneuron morphology observed in *Dlx1* null mutants supports this contention[28]. These findings indicate that genetic programs initiated at the progenitor stage are modulated during development by activity. Thus, our studies indicate that the role of early network activity in shaping the development of specific neuronal subtypes in the central nervous system is greater than is presently appreciated.

METHODS SUMMARY

Mouse strains and *in utero* electroporation. Pregnant wild-type and genetically modified mice (see Methods) were electroporated at 15 days of gestation (E15.5) using a standard *in utero* electroporation technique[31]. The plasmids used in the electroporation experiments were generated using standard cloning techniques.
***In situ* hybridization and immunohistochemistry.** *In situ* hybridization and immunohistochemistry were performed as previously described[32]. For morphological reconstruction, vibratome sections were fixed and incubated overnight at 4 °C with selected antibodies.
Quantification of interneuron layer distribution. The proportion of Cr^{+}, Re^{+} and VIP^{+} interneurons over the total number of electroporated interneurons across cortical layers was calculated in all cryostat tissue sections from individual brains. Tbr1 immunolabelling was used to delineate cortical layers II/III and V at P5–P8.
Kynurenic acid treatment. *Dlx5/6-eGFP* electroporated pups were anaesthetized by hypothermia. Kynurenic acid (300 nM, Sigma-Aldrich) diluted in PBS or pure PBS (controls) were injected at P0, P2 and P3. Treated brains in which electroporated interneurons were found in the vicinity of the injection site were used for analysis to minimize variability due to drug diffusion.
Electrophysiology. Whole-cell patch-clamp electrophysiological recordings were performed on eGFP-expressing cells in acute brain slices prepared from P8–P18 animals. Whole-cell recordings were made from randomly selected eGFP-positive neurons located in the upper layers (I–III) of the somatosensory cortex. Experiments were performed in both current-clamp and voltage-clamp modes.
Neuronal morphology analysis. Images of interneurons were obtained with a confocal microscope, analysed with LSM Image Browser, and reconstructed with Neurolucida software (Version 9). To assess the length and complexity of dendritic and axonal arborizations, we quantified the number of nodes (points from which two or more branches arose) and ends (terminal branches) in each of these trees with Neurolucida Explorer.
Statistical analysis. Statistical analysis was performed by using Student's *t*-test (two-tailed distribution, homoscedastic) unless otherwise stated.

Detailed methods on the mouse strains, animal surgery and electrophysiology protocols can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 July 2010; accepted 18 January 2011.
Published online 3 April; corrected 21 April 2011 (see full-text HTML version for details).

1. Blankenship, A. G. & Feller, M. B. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nature Rev. Neurosci.* **11,** 18–29 (2010).
2. Wong, R. O., Chernjavsky, A., Smith, S. J. & Shatz, C. J. Early functional neural networks in the developing retina. *Nature* **374,** 716–718 (1995).
3. Penn, A. A., Riquelme, P. A., Feller, M. B. & Shatz, C. J. Competition in retinogeniculate patterning driven by spontaneous activity. *Science* **279,** 2108–2112 (1998).
4. Huberman, A. D. *et al.* Architecture and activity-mediated refinement of axonal projections from a mosaic of genetically identified retinal ganglion cells. *Neuron* **59,** 425–438 (2008).
5. Spitzer, N. C. Electrical activity in early neuronal development. *Nature* **444,** 707–712 (2006).
6. Root, C. M., Velazquez-Ulloa, N. A., Monsalve, G. C., Minakova, E. & Spitzer, N. C. Embryonically expressed GABA and glutamate drive electrical activity regulating neurotransmitter specification. *J. Neurosci.* **28,** 4777–4784 (2008).
7. Cancedda, L., Fiumelli, H., Chen, K. & Poo, M. M. Excitatory GABA action is essential for morphological maturation of cortical neurons *in vivo*. *J. Neurosci.* **27,** 5224–5235 (2007).
8. Wang, C. L. *et al.* Activity-dependent development of callosal projections in the somatosensory cortex. *J. Neurosci.* **27,** 11334–11342 (2007).
9. Bortone, D. & Polleux, F. KCC2 expression promotes the termination of cortical interneuron migration in a voltage-sensitive calcium-dependent manner. *Neuron* **62,** 53–71 (2009).
10. Ascoli, G. A. *et al.* Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nature Rev. Neurosci.* **9,** 557–568 (2008).
11. Batista-Brito, R. & Fishell, G. The developmental integration of cortical interneurons into a functional network. *Curr. Top. Dev. Biol.* **87,** 81–118 (2009).
12. Miyoshi, G. *et al.* Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *J. Neurosci.* **30,** 1582–1594 (2010).
13. Karube, F., Kubota, Y. & Kawaguchi, Y. Axon branching and synaptic bouton phenotypes in GABAergic nonpyramidal cell subtypes. *J. Neurosci.* **24,** 2853–2865 (2004).
14. Gumienny, T. L. *et al.* CED-12/ELMO, a novel member of the CrkII/Dock180/Rac pathway, is required for phagocytosis and cell migration. *Cell* **107,** 27–41 (2001).
15. Cobos, I., Borello, U. & Rubenstein, J. L. Dlx transcription factors promote migration through repression of axon and dendrite growth. *Neuron* **54,** 873–888 (2007).
16. Allene, C. & Cossart, R. Early NMDA receptor-driven waves of activity in the developing neocortex: physiological or pathological network oscillations? *J. Physiol. (Lond.)* **588,** 83–91 (2010).
17. Garaschuk, O., Linn, J., Eilers, J. & Konnerth, A. Large-scale oscillatory calcium waves in the immature cortex. *Nature Neurosci.* **3,** 452–459 (2000).
18. Dupont, E., Hanganu, I. L., Kilb, W., Hirsch, S. & Luhmann, H. J. Rapid developmental switch in the mechanisms driving early cortical columnar networks. *Nature* **439,** 79–83 (2006).
19. Stenman, J., Toresson, H. & Campbell, K. Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. *J. Neurosci.* **23,** 167–174 (2003).
20. Yu, C. R. *et al.* Spontaneous neural activity is required for the establishment and maintenance of the olfactory sensory map. *Neuron* **42,** 553–566 (2004).
21. Yang, J. W., Hanganu-Opatz, I. L., Sun, J. J. & Luhmann, H. J. Three patterns of oscillatory activity differentially synchronize developing neocortical networks *in vivo*. *J. Neurosci.* **29,** 9011–9025 (2009).
22. Khazipov, R. & Luhmann, H. J. Early patterns of electrical activity in the developing cerebral cortex of humans and rodents. *Trends Neurosci.* **29,** 414–418 (2006).
23. McCabe, A. K., Chisholm, S. L., Picken-Bahrey, H. L. & Moody, W. J. The self-regulating nature of spontaneous synchronized activity in developing mouse cortical neurones. *J. Physiol. (Lond.)* **577,** 155–167 (2006).
24. Manent, J. B., Jorquera, I., Ben-Ari, Y., Aniksztejn, L. & Represa, A. Glutamate acting on AMPA but not NMDA receptors modulates the migration of hippocampal interneurons. *J. Neurosci.* **26,** 5901–5909 (2006).
25. Stone, T. W. Neuropharmacology of quinolinic and kynurenic acids. *Pharmacol. Rev.* **45,** 309–379 (1993).
26. Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. Interneuron migration from basal forebrain to neocortex: dependence on *Dlx* genes. *Science* **278,** 474–476 (1997).
27. Wonders, C. P. & Anderson, S. A. The origin and specification of cortical interneurons. *Nature Rev. Neurosci.* **7,** 687–696 (2006).
28. Cobos, I. *et al.* Mice lacking *Dlx1* show subtype-specific loss of interneurons, reduced inhibition and epilepsy. *Nature Neurosci.* **8,** 1059–1068 (2005).
29. Ravichandran, K. S. & Lorenz, U. Engulfment of apoptotic cells: signals for a good meal. *Nature Rev. Immunol.* **7,** 964–974 (2007).
30. Park, D. *et al.* BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. *Nature* **450,** 430–434 (2007).
31. Saito, T. *In vivo* electroporation in the embryonic mouse central nervous system. *Nature Protocols* **1,** 1552–1558 (2006).
32. Butt, S. J. *et al.* The requirement of *Nkx2-1* in the temporal specification of cortical interneuron subtypes. *Neuron* **59,** 722–732 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to R. Batista-Brito, E. Chiappe, R. Cossart, J. Dasen, J. Kaltschmidt, S. Lee, J. Hjerling Leffler, M. Long, D. Pisapia and B. Rudy for comments on the manuscript. We thank L. Yin for technical assistance. We are indebted to K. Ravichandran for providing the ELMO1 constructs. N.V.D.G. and T.K. are both supported by grants from The Patterson Trust. Research in the Fishell laboratory is supported by the National Institutes of Health, National Institute of Mental Health (5RO1MH068469-08 and 2R01MH071679-09), National Institute of Neurological Disorders and Stroke (5R01NS039007-1), New York Stem Cell Science State (NGSG-130) and the Simons Foundation.

Author Contributions N.V.D.G. and G.F. conceived the project. N.V.D.G. and T.K. designed and carried out the experiments. N.V.D.G. wrote the manuscript with the help of all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.F. (fisheg01@nyumc.org).

METHODS

Mouse strains. Pregnant Swiss Webster mice (Taconic) were electroporated at 15 days of gestation (E15.5). The *tetO-Kir2.1.ires.LacZ* transgenic mouse line was provided by J. Gogos[20]. Doxycycline was administered in mouse feed (20 g per kg of feed) at selected time points (E16.5, P0, P3). The *Gad67-GFP* (gift from Y. Yanagawa) mouse line[33] was available in the G.F. laboratory. Details on the genotyping of the mouse strains have been described elsewhere[12].

***In utero* electroporation.** Pregnant mice were electroporated using a standard *in utero* electroporation technique[31]. In brief, a timed pregnant mouse was anaesthetized and embryos were injected through the uterine wall in one lateral ventricle with 1–2 μl of DNA (3 $\mu g\,\mu l^{-1}$). Fast green was used for visualization of the DNA solution. DNA was delivered by a glass needle operated with a mouth pipette. Five square 50-ms pulses at 40 mV with a 950 ms interval were delivered with a 5-mm paddle electrode (CUY650P5, Protech Iternational) using an electroporator (CUY21, Protech International). After electroporation, the uterus was placed back in the abdominal cavity and the mouse was sutured. The mice were kept on a warm plate (Fine Science Tools) through surgery to minimize hypothermia. After surgery, mice recovered in a humidified chamber at 30 °C for 2–3 h. Mouse colony maintenance and handling was performed in compliance with the protocols approved by the Institutional Animal Care and Use Committee of the New York University School of Medicine.

The plasmids used in the electroporation experiments were generated using standard cloning techniques. The mouse *Kir2.1*, *Tta*, *mCherry*, *eGFP*, *Elmo1* and *Elmo1.TN558.Flag* cDNAs were each individually cloned into a *Dlx5/6-Pmin-polyA* plasmid. Because eGFP expression was not detected in brains electroporated with a *Dlx-5/6-Kir2.1.ires.eGFP* polycistronic plasmid, the *Dlx5/6-eGFP* plasmid was co-electroporated with *Dlx5/6-Kir2.1*, *Dlx5/6-Tta*, *Dlx5/6-Elmo1* or *Dlx5/6-Elmo1.TN558.Flag* plasmids at equivalent molar concentrations to ensure high levels of co-expression. The detection of similar levels of eGFP expression in *Dlx5/6-eGFP* and *Dlx5/6-eGFP/Dlx5/6-Kir2.1* electroporated interneurons indicates that transcription driven by this enhancer is not affected by Kir2.1 expression. For generation of *CAG-mCherry*, the *mCherry* cDNA was cloned into a *CAG-MCS* vector. Expression of the *tetO-Kir2.1.ires.LacZ* transgene in interneurons electroporated with the *Dlx5/6-Tta* plasmid was detected by processing tissue sections for β-galactosidase staining[20].

***In situ* hybridization and immunohistochemistry.** *In situ* hybridization was performed as described[32] using a full-length *Kir2.1* dig-labelled probe. Immunohistochemistry on 20-μm tissue cryostat sections was previously described[34]. For morphological reconstruction, 250-μm-thick vibratome sections were fixed for 2 h and incubated overnight at 4 °C with selected antibodies. Sections were washed in PBS for several hours and incubated at 4 °C overnight with donkey secondary antibodies (Jackson laboratories). Primary antibodies used in the experiments include rat anti-GFP (1:2,000; Nacalai Tesque), mouse anti-Reelin (CR50) (1:500; MBL), rabbit anti-VIP (1:1,000; Immunostar), mouse anti-calretinin (1:1,500; Millipore Bioscience Research Reagents), rabbit anti-Tbr1 (1:1,000; Abcam), goat anti-Tbr1 (1:1,000; Abcam), rabbit anti-Pan-DLX (a gift from J. Kohtz), rabbit anti-NPAS1 (a gift from M. Masayuki), goat anti-ELMO1 (1:250; Millipore) and mouse anti-Flag (1:200; Sigma-Aldrich).

Quantification of cell death. Caspase 3 activity (Clontech) was assessed on cryostat sections of P8 brains electroporated with *Dlx5/6-eGFP* or *Dlx5/6-eGFP* and *Dlx5/6-Kir2.1* plasmids. The percentage of Caspase3 immunoreactive interneurons that co-express eGFP over the total number of eGFP-expressing interneurons was counted on five Kir2.1-electroporated mice and five control mice.

Quantification of interneuron layer distribution. The proportion of Cr^{+}, Re^{+} and VIP^{+} interneurons over the total number of electroporated interneurons across cortical layers was calculated in all cryostat tissue sections from individual brains. Analysis was performed on four wild-type (74 interneurons) and six *tetO-Kir2.1.ires.LacZ* (150 interneurons) mice co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids.

Kynurenic acid treatment. *Dlx5/6-eGFP* electroporated pups were anaesthetized by hypothermia on ice for two minutes. The pups were protected with cloth to prevent frostbite. Kynurenic acid[25] (300 nM, Sigma-Aldrich) diluted in PBS and pure PBS (controls) was injected at P0, P2 and P3. Fast green was used for visualization. A small window was opened in the skull with needles and solution was injected in the subdural space on the electroporated side. The skull opening was closed with cynoacrylate adhesive. Pups were allowed to recover in a humidified chamber at 34 °C for 5–10 min and another 20 min at room temperature (18 °C) before putting them back in their cages. Treated brains in which electroporated interneurons were found in the vicinity of the injection site were used for analysis to minimize variability due to drug diffusion. Kynurenic acid injections at P0, P1, P2 and P3 had no effect on interneuron migration. Therefore, we averaged the values obtained for laminar distribution in six control (110 interneurons) and six kynurenic-acid-treated (165 interneurons) mice electroporated with a *Dlx5/6-eGFP* plasmid. Analysis was performed at P8–P9.

Electrophysiology. Whole-cell patch-clamp electrophysiological recordings were performed on eGFP-expressing cells in acute brain slices prepared from P8–P18 animals. Briefly, animals were decapitated and the brain was dissected out and transferred to physiological artificial cerebrospinal fluid (ACSF) cooled down to 4 °C of the following composition: 125 mM NaCl, 2.5 mM KCl, 25 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 1 mM $MgCl_2$, 2 mM $CaCl_2$ and 20 mM glucose. The brain was then glued to a stage and 250-μm-thick slices were cut using a vibratome (Vibratome 3000 EP). The slices were allowed to recover in recording ACSF at room temperature for at least 45 min before recording. They were then placed in a recording chamber mounted on the stage of an upright microscope (Axioscope, Zeiss) equipped with immersion differential interference contrast objectives (×5, ×40) coupled to an infrared camera system (Zeiss), superfused at a rate of 1–2 ml min^{-1} with oxygenated recording ACSF and maintained at a temperature of 31 °C. An eGFP filter was used to visualize the fluorescent interneurons in epifluorescence.

Whole-cell recordings were made from randomly selected eGFP-positive neurons located in upper layers (I–III) of the somatosensory cortex. Patch electrodes were made from borosilicate glass (Harvard Apparatus), had a resistance of 4–8 M and were filled with a solution containing: 128 mM K-gluconate, 4 mM NaCl, 0.3 mM Na-GTP, 5 mM Mg-ATP, 0.0001 mM $CaCl_2$, 10 mM HEPES, 1 mM glucose and 5 mg ml^{-1} biocytin (Sigma). Experiments were performed in current-clamp mode using the Axoclamp 2B (Molecular Devices) or the Axopatch 200B amplifier and in voltage clamp using the latter.

Access resistance was always monitored to ensure the stability of recording conditions. Cells were only accepted for analysis if the initial series resistance was less than or equal to 40 MΩ and did not change by more than 20% throughout the recording period. The series resistance was compensated online by at least ~ 50% in voltage-clamp mode to reduce voltage errors. No correction was made for the junction potential between the pipette and the ACSF.

For Kir2.1 conductance assessment a series of voltage steps in 10-mV increments were applied every 1–5 s in voltage clamp from −140 mV to 0 mV starting from −70 mV after a prepulse down to −90 mV so as to deinactivate any Kir2.1 channels that had entered inactivated states.

Active firing and passive membrane properties were recorded in current-clamp mode by applying a series of sub- and suprathreshold current steps. The resting membrane potential (V_{rest}) was ascertained in current clamp right after rupturing the patch by applying zero current.

All drugs were applied to the recording preparation through the bath. Salts used in the preparation of the intracellular recording solution and ACSF were obtained from Sigma-Aldrich. Kynurenic acid and SR95531 were also purchased from Sigma-Aldrich.

Neuronal morphology analysis. Images of interneurons were obtained with a Zeiss (LSM 510 Meta) confocal microscope, analysed with LSM Image Browser, and reconstructed with Neurolucida software (v. 9). Morphological defects were observed in >50 interneurons (>10 brains) of each subtype (Cr^{+} and Re^{+}) after Kir2.1 electroporation and kynurenic acid treatment. In addition, analysis of morphology after doxycycline administration in *tetO-Kir2.1.ires.LacZ* mice co-electroporated with *Dlx5/6-Tta* and *Dlx5/6-eGFP* plasmids was performed in >20 interneurons (>4 brains). Similarly, >70 interneurons (>10 brains) were analysed in control experiments. A few of these interneurons were chosen for reconstruction. The total length and complexity of axonal arborizations and dendritic trees was scored in confocal stacks (optical slice thickness, 4 μm; stack size 50–100 μm) including all the neuronal processes. Interneurons are oriented such that the top of the figure panel points towards the pia and the bottom to the lateral ventricle. To assess the length and complexity of dendritic and axonal arborizations, we quantified the number of nodes (points from which two or more branches arose) and ends (terminal branches) in each of these trees with Neurolucida Explorer. Total length and complexity of neuronal processes were scored in the same set of reconstructed interneurons for each experiment. eGFP labelling in electroporated interneuron was indistinguishable form that of streptavidin fills.

Statistical analysis. Statistical analysis was performed by using Student's *t*-test (two-tailed distribution, homoscedastic) unless otherwise stated.

33. Tamamaki, N. *et al.* Green fluorescent protein expression and colocalization with calretinin, parvalbumin, and somatostatin in the GAD67-GFP knock-in mouse. *J. Comp. Neurol.* **467,** 60–79 (2003).
34. Miyoshi, G., Butt, S. J., Takebayashi, H. & Fishell, G. Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. *J. Neurosci.* **27,** 7786–7798 (2007).

LETTER

doi:10.1038/nature09874

Ephrin Bs are essential components of the Reelin pathway to regulate neuronal migration

Aycan Sentürk[1], Sylvia Pfennig[1], Alexander Weiss[1]†, Katja Burk[1]† & Amparo Acker-Palmer[1]

Coordinated migration of neurons in the developing and adult brain is essential for its proper function. The secreted glycoprotein Reelin (also known as RELN) guides migration of neurons by binding to two lipoprotein receptors, the very-low-density lipoprotein receptor (VLDLR) and apolipoprotein E receptor 2 (ApoER2, also known as LRP8)[1]. Loss of Reelin function in humans results in the severe developmental disorder lissencephaly[2] and it has also been associated with other neurological disorders such as epilepsy, schizophrenia and Alzheimer's disease[3]. The molecular mechanisms by which Reelin activates its receptors and controls cellular functions are largely unknown. Here we show that the neuronal guidance cues ephrin B proteins are essential for Reelin signalling during the development of laminated structures in the brain. We show that ephrin Bs genetically interact with Reelin. Notably, compound mouse mutants (*Reln*$^{+/-}$; *Efnb3*$^{-/-}$ or *Reln*$^{+/-}$; *Efnb2*$^{-/-}$) and triple ephrin B1, B2, B3 knockouts show neuronal migration defects that recapitulate the ones observed in the neocortex, hippocampus and cerebellum of the *reeler* mouse. Mechanistically, we show that Reelin binds to the extracellular domain of ephrin Bs, which associate at the membrane with VLDLR and ApoER2 in neurons. Clustering of ephrin Bs leads to the recruitment and phosphorylation of Dab1 which is necessary for Reelin signalling. Conversely, loss of function of ephrin Bs severely impairs Reelin-induced Dab1 phosphorylation. Importantly, activation of ephrin Bs can rescue the *reeler* neuronal migration defects in the absence of Reelin protein. Together, our results identify ephrin Bs as essential components of the Reelin receptor/signalling pathway to control neuronal migration during the development of the nervous system.

The mammalian neocortex has characteristic laminations, containing different neuron types arranged in stereotypical patterns. Formation of this layered structure is possible only with proper migration of neurons from proliferative zones to their final position[4–6]. In recent years the extracellular protein Reelin has emerged as an important factor that affects several steps of neuronal migration and layering in the cerebral cortex (reviewed in ref. 7). Reelin exerts its function by binding to the lipoprotein receptors VLDLR and ApoER2 and inducing the phosphorylation of the adaptor protein Dab1 (refs 1, 8) by Src-family kinases (SFKs)[9,10]. Despite the importance of the Reelin signalling pathway for proper nervous system development and that disruption of this pathway results in the severe developmental disorder lissencephaly[2] and is associated with epilepsy, schizophrenia and Alzheimer's disease[3,11,12], molecular characterization of the activation of this pathway at the cell membrane remains poorly understood. Because VLDLR and ApoER2 do not possess intrinsic kinase activity the existence of a co-receptor that locally activates Src kinases has been proposed for at least a decade. We have previously shown that ephrin Bs, transmembrane ligands for Eph receptors, have signalling capabilities that are required for synaptic plasticity and sprouting angiogenesis by regulating the activity of other transmembrane receptors such as α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor and vascular endothelial growth factor (VEGF) receptor 2, respectively[13–15]. Moreover, stimulation of cultured cortical neurons with soluble Eph B receptors leads to the recruitment and activation of SFKs in ephrin B-membrane patches[16]. Therefore, we proposed that ephrin Bs might control Reelin signalling *in vivo*.

To study whether ephrin B and Reelin signalling genetically interact *in vivo* we generated compound mutant mice null for ephrin B3 or ephrin B2 and heterozygous for Reelin (*Reln*$^{+/-}$; *Efnb3*$^{-/-}$ and *Reln*$^{+/-}$; *Efnb2*$^{-/-}$). *Reeler* is an autosomal recessive mutant mouse and in heterozygosity (*Reln*$^{+/-}$) does not show any overt phenotypes[17]. Gross morphology of the cortex assessed by DAPI staining on coronal cortical sections showed alterations in cell layering in the *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice compared to control littermates (Supplementary Fig. 1a). NeuN staining revealed a small but significant increase in neurons in the marginal zone of the ephrin B3 compound mutant mice (*Reln*$^{+/-}$; *Efnb3*$^{-/-}$), one of the hallmarks of the *reeler* phenotype[18] (Supplementary Fig. 1b, c). To characterize the cortical defects in detail we analysed the distribution of postmitotic migrating neurons in the cortex of *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mutant mice using different layer-specific markers. In wild-type mice, Tbr1 is expressed in some cortical plate neurons in layer II and most abundantly in neurons in layer VI (Fig. 1a, b) and therefore is a typical marker for early born glutamergic neocortical neurons (lower layers)[19]. Analysis of the *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compounds revealed the presence of abundant Tbr1-positive (Tbr1^{+}) neurons instead in the superficial layers of the cortex (Fig. 1a, b). We confirmed the aberrant presence of large pyramidal neurons expressing green fluorescent protein (GFP) in upper cortical layers by crossing the *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mutant or the *reeler* mice with Thy1–GFP M-line (Supplementary Fig. 2). Conversely, markers for late cortical plate neurons (upper layers II–IV), Brn1 (also known as POU3F3) and SatB2 (ref. 20) showed an aberrant distribution in *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice with an accumulation of late born neurons in the lower cortical layers (Fig. 1c and Supplementary Fig. 3). At embryonic day 17.5 (E17.5), *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice mimicked the aberrant cortical plate splitting of the *reeler* mice[21] with a striking accumulation of the extracellular matrix component chondroitin sulphate proteoglycan (CSPG) throughout the neocortex (Supplementary Fig. 4). To confirm that the observed defects in cortex lamination in the *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice are due to neuronal migration impairment we performed additional 5-bromodeoxyuridine (BrdU) pulse experiments (Supplementary Fig. 5). BrdU is incorporated in dividing progenitor cells and therefore reflects the migratory behaviour of newly born neurons at the time of injection. Quantification of cell distribution in the different cortical layers showed the characteristic *reeler* inverted cortical layering in the ephrin B3 compound mice (Supplementary Fig. 5). Loss of ephrin Bs does not cause defects in differentiation of Reelin-positive Cajal–Retzius cells or the level of expression of Reelin and components of the Reelin pathway (Supplementary Fig. 6). Therefore, all these results confirm a *reeler* outside-in cortex layering in

[1]Frankfurt Institute for Molecular Life Sciences (FMLS) and Institute of Cell Biology and Neuroscience, Goethe University Frankfurt, Max-von-Laue-Str. 9, D-60438, Frankfurt am Main, Germany. †Present addresses: Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Room 1078, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada (A.W.); IBDML-UMR6216-CNRS, Case 907- Parc Scientifique de Luminy, 13288 Marseille Cedex 9, France (K.B.).

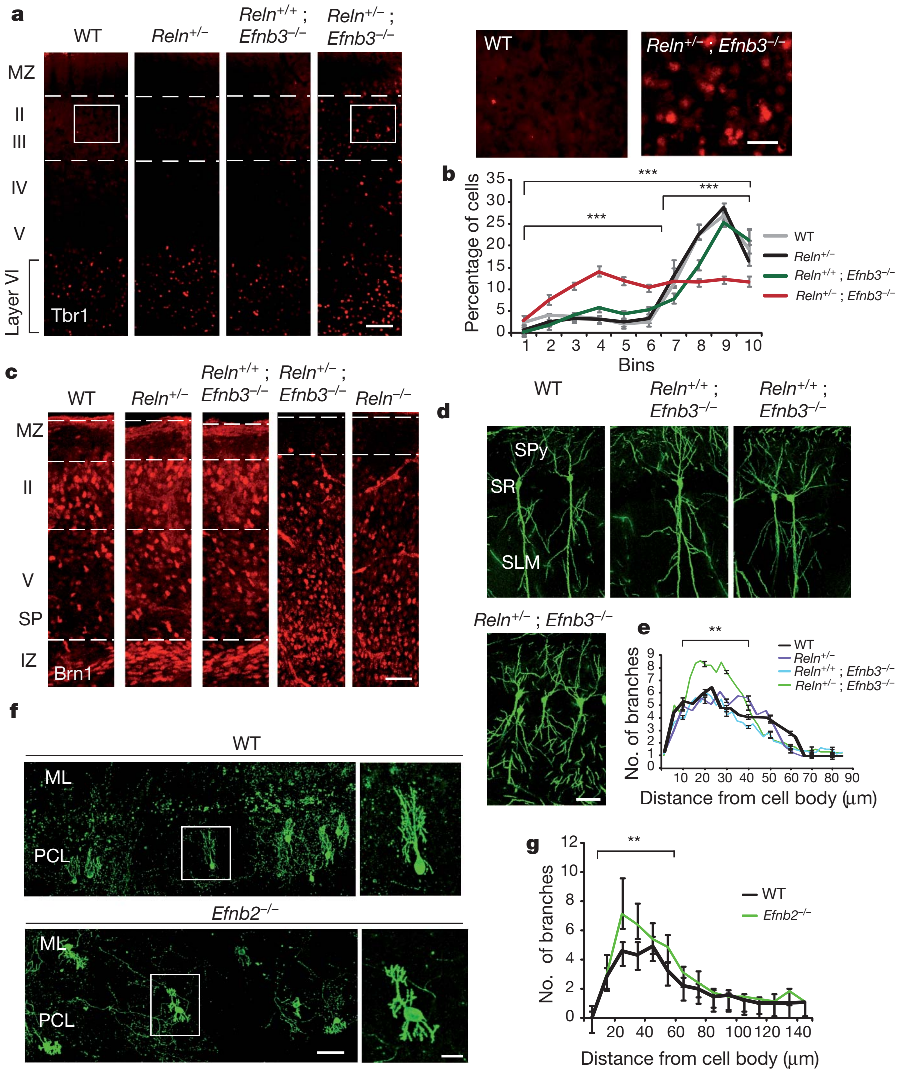


Figure 1 | Ephrin B3 and ephrin B2 deficiency differentially phenocopy *reeler*-like defects. **a**, Tbr1 staining of cerebral cortex of adult mice. Tbr1^{+} cells accumulated in the upper cortical layers (II–III, dashed lines) of the compound mice (*Reln*$^{+/-}$; *Efnb3*$^{-/-}$) compared to wild-type (WT), heterozygous *reeler* (*Reln*$^{+/-}$) and ephrin B3 knockout (*Efnb3*$^{-/-}$) mice. Magnification pictures show accumulation of cells in layer II–III of *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice compared to control littermates (WT). **b**, Quantification based on the distribution of the Tbr1^{+} cells in the cerebral cortex (s.e.m., $n = 10$–30). **c**, Immunostaining for Brn1 of E17.5 mice cortex showed late born neurons accumulated in lower layers in ephrin B3 compound mice similar to *reeler* cortex. **d**, Pyramidal neurons projecting towards the stratum locunosum moleculare (SLM) show misorientation and branching defects in the ephrin B3 compound mice (*Reln*$^{+/-}$; *Efnb3*$^{-/-}$, Thy1–GFP^{+}). **e**, Sholl analysis showed increased branching close to the cell body (s.e.m., $n = 90$–120). **f**, **g**, Dendrite structure of Purkinje cells in the Purkinje cell layer was visualized by GFP electroporation of P6 organotypic cerebellar cultures and analysed by Sholl analysis. Increased arborization was observed in the single ephrin B2 knockout mice (*Efnb2*$^{-/-}$) (s.e.m., $n = 30$–40). Scale bars, 100 μm (**f**); 75 μm (**a**, **c**); 50 μm (higher magnifications in **f**), 25 μm (**d**, higher magnifications of **a**). IZ, intermediate zone; ML, molecular layer; MZ, marginal zone; PCL, Purkinje cell layer; SLM, stratum lacunosum moleculare; SP, sub-plate, SPy, stratum pyramidale; SR, stratum radiatum. $**P < 0.01$; $***P < 0.001$.

the *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice and suggest that these lamination defects are in fact a result of an impaired neuronal migration.

Hippocampal laminated structures were also found to be abnormal in *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice. In these mice, the CA1 region of the hippocampus showed significant increase in thickness and invasion of pyramidal cells into the stratum oriens (Supplementary Fig. 7a, b). Moreover, the abnormalities observed in *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice in the dendritic branching of pyramidal neurons of the CA1 region strikingly resembled the ones observed in the *reeler* mutants (Fig. 1d, e and Supplementary Fig. 8). Reelin also controls differentiation and position of GFAP^{+} radial glial cells for the proper migration of newly generated granule cells[22]. In *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice this radial glial scaffold network was severely affected (Supplementary Fig. 9) and this also resulted in a loose organization of the granule cell layers in the dentate gyrus (Supplementary Fig. 10). We have also observed defects in the cortex and hippocampus of compound *Reln*$^{+/-}$; *Efnb2*$^{-/-}$ mice (data not shown) indicating that both ephrin B2 and ephrin B3 might be involved and be redundant in their function to control Reelin signalling during neuronal migration in both brain structures.

Ataxia in *reeler* mice results from a severe disorganization of the cerebella in those mutants[7]. Interestingly, in *Reln*$^{+/-}$; *Efnb3*$^{-/-}$ compound mice we found only mild defects in the cerebellar structures (Supplementary Fig. 11a–d, g). Instead, ephrin B2 is highly expressed in the Purkinje cells in the cerebellum[23] and although no major defects on the migration of these cells were observed (Supplementary Fig. 11h), *Efnb2*$^{-/-}$ mice showed striking defects in the arborization of Purkinje cells analysed by electroporation of GFP in acute cerebellar slices (Fig. 1f, g) as well as by using calbindin staining (Supplementary Fig. 11e, f). Importantly, arborization defects that were already visible in the single ephrin B2 mutants (*Efnb2*$^{-/-}$) were significantly increased in the *Reln*$^{+/-}$; *Efnb2*$^{-/-}$ compound mice (Supplementary Fig. 11e, f),

indicating that, in the cerebellum, Reelin functions in the arborization of Purkinje cells require specifically ephrin B2 and not ephrin B3.

Analysis of the compound mice revealed that ephrin B2 and ephrin B3 might have compensatory, but also specific functions in controlling Reelin signalling in the different brain structures. Importantly, a triple ephrin B1, B2, B3 mutant displayed severe migration phenotypes that resembled the *reeler* mouse (Fig. 2a–c and Supplementary Fig. 12), indicating a major role for ephrin Bs in the Reelin signalling pathway *in vivo*. However, some phenotypes in the cerebellum and hippocampus of the ephrin B1, B2, B3 mutant are not as strong as in the *reeler* mouse, suggesting compensation by ephrin B-independent functions of Reelin. In agreement with an important role of ephrin Bs controlling Reelin functions in the developing brain we find ephrin Bs colocalizing with Reelin receptors and Dab1 in migrating neurons in the cerebral cortex (Fig. 2d, e), hippocampus (Supplementary Fig. 13a, b) and in Purkinje cell neurons in the developing cerebellum (Supplementary Fig. 13c).

In order to get mechanistic insights on how ephrin Bs can crosstalk to Reelin signalling we next assessed the biochemical interaction of these two pathways. We performed a directed proteomic analysis of ephrin B-binding proteins from a neuroblastoma cell line using the tandem affinity purification–mass spectrometry methodology[24] and identified Reelin as a putative ephrin B-interacting protein (data not shown). We first confirmed the ability of Reelin to associate with ephrin Bs by co-immunoprecipitation of both endogenous proteins from brain lysates (Fig. 3a). Reelin extracellular protein bound directly to the extracellular domain of ephrin B3 and also to ephrin B2 (Fig. 3b), suggesting that both ephrins Bs could influence the *in vivo* functions of Reelin signalling. Importantly, stimulation of cortical neurons with Reelin leads to the clustering of ephrin Bs (Fig. 3c) and effective tyrosine phosphorylation of the adaptor Dab1 (Supplementary Fig. 14). We have previously shown that activation of ephrin Bs leads to the recruitment and activation of Src kinases in ephrin B clusters[16]. Therefore, we propose that ephrin Bs could recruit and activate Src kinases in VLDLR and ApoER2 receptor clusters. In agreement, we find that phosphorylated Dab1 co-immunoprecipitates with ephrin Bs at E16.5 when the neuronal migration and layering of the brain structures takes place (Fig. 3d and Supplementary Fig. 15a). Activation of ephrin Bs by a soluble receptor Eph B3–Fc leads to the recruitment and phosphorylation of Dab1 in ephrin B clusters in cortical and hippocampal neurons (Fig. 3e and Supplementary Fig. 15b). Importantly, we find that ephrin Bs co-immunoprecipitate with both Reelin receptors, ApoER2 and VLDLR, (Fig. 3f, g) and activation of ephrin B by Eph B3–Fc clusters these receptors in ephrin B3 membrane patches in cortical and hippocampal neurons (Supplementary Fig. 15c–f) without affecting their level of expression (Supplementary Fig. 16). Co-stimulation of ephrin Bs by Eph B receptors and Reelin in cortical neurons resulted in increased Dab1 phosphorylation (Supplementary Fig. 17a–c). This function of ephrin Bs requires Src kinases (Supplementary Fig. 17d) and fully functional Reelin receptors because VLDLR knockouts and ApoER2 knockouts both showed a reduction in the activation of Dab1 following Eph B3–Fc treatment in cortical neurons (Fig. 3h and Supplementary Fig. 17e–h).

We next addressed the requirement of ephrin B-mediated recruitment and activation of Src kinases for Reelin signalling by loss-of function studies both in cortical neurons in culture and *in vivo* in mice. Cortical neurons isolated from ephrin B3 and ephrin B2 double knockouts showed a significant impairment of Reelin-mediated phosphorylation of Dab1 (Fig. 4a, b). Ephrin B3 mouse mutants also showed *in vivo* a reduction in Dab1 phosphorylation that correlated with increased levels of Dab1 proteins (Fig. 4c–e) as it has been already shown for different mouse mutants in Reelin signalling components[8,9,25]. More importantly, the levels of phosphorylation of Dab1 in triple ephrin B1, B2, B3 mouse mutants were greatly reduced to the same extent as in *reeler* mutants (Fig. 4c–e). We next set up a series of rescue experiments both in neuronal cultures as well as during neuronal migration in the cortex *in vivo*. Phosphorylation of Dab1 in cortical neurons isolated from *reeler* mice could be rescued as expected by stimulation with exogenous Reelin (Fig. 4f and Supplementary Fig. 18). More importantly, Dab1 phosphorylation in *reeler* neurons was rescued solely by activation of ephrin Bs with Eph receptors. To underline the importance of ephrin B for Reelin signalling during neuronal migration in the cerebral cortex we set up similar rescue experiments of the *reeler* phenotype in organotypic slice cultures. *Reeler* slices showed the typical *reeler* phenotype in the cerebral cortex with BrdU^{+} cells accumulated strongly in the lower layers. Importantly, ephrin B activation by Eph receptors was able to rescue the migratory defects in *reeler* slices guiding the BrdU^{+} cells to their proper location on layers II–III (Fig. 4g, h).

Taken together, our genetic analysis together with a strong biochemical analysis of the ephrin B mutant mice identify ephrin Bs as

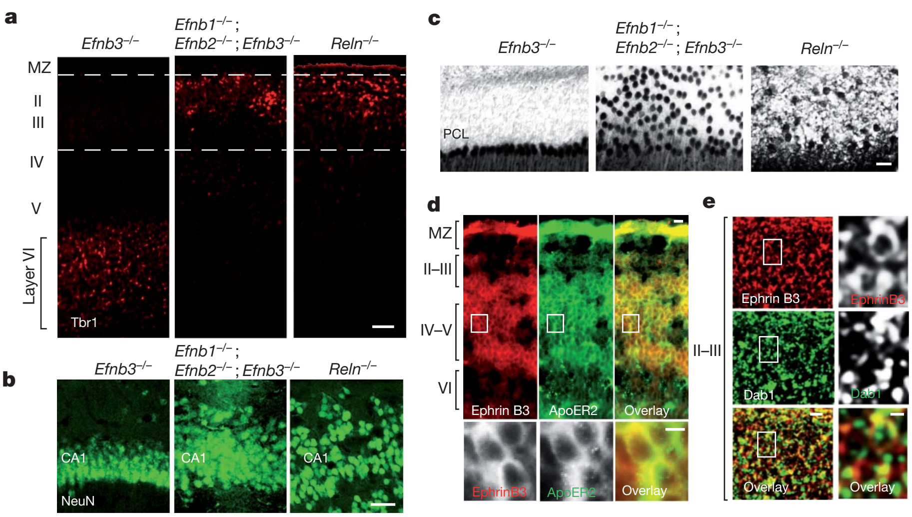


Figure 2 | Ephrin B1, B2, B3 triple mutants show a *reeler*-like phenotype. **a**, Cortical layering is inverted in the triple ephrin B1, B2, B3 mice shown in coronal cortical sections of P2 mice immunostained for Tbr1 for early born neurons (layers defined with dashed lines). **b**, NeuN analysis of adult hippocampus of *Efnb3*$^{-/-}$, *Efnb1*$^{-/-}$; *Efnb2*$^{-/-}$; *Efnb3*$^{-/-}$ and *Reln*$^{-/-}$ mice showing an expansion of the CA1 region. **c**, Calbindin immunostaining of P20 cerebellum of *Efnb3*$^{-/-}$, *Efnb1*$^{-/-}$; *Efnb2*$^{-/-}$; *Efnb3*$^{-/-}$ and *Reln*$^{-/-}$ showed severe migration defects in the Purkinje cell layer (PCL) in the *Efnb1*$^{-/-}$; *Efnb2*$^{-/-}$; *Efnb3*$^{-/-}$. **d**, **e**, Colocalization of ephrin B3 with ApoER2 and Dab1 detected by immunostaining in the cerebral cortex of E17.5 mice. Scale bars: 75 μm (**d**, **e**), 50 μm (**a**, **b**, **c**), 25 μm (higher magnifications of **d**, **e**). CA1, Hippocampal area; MZ, Marginal zone, PCL, Purkinje cell layer.

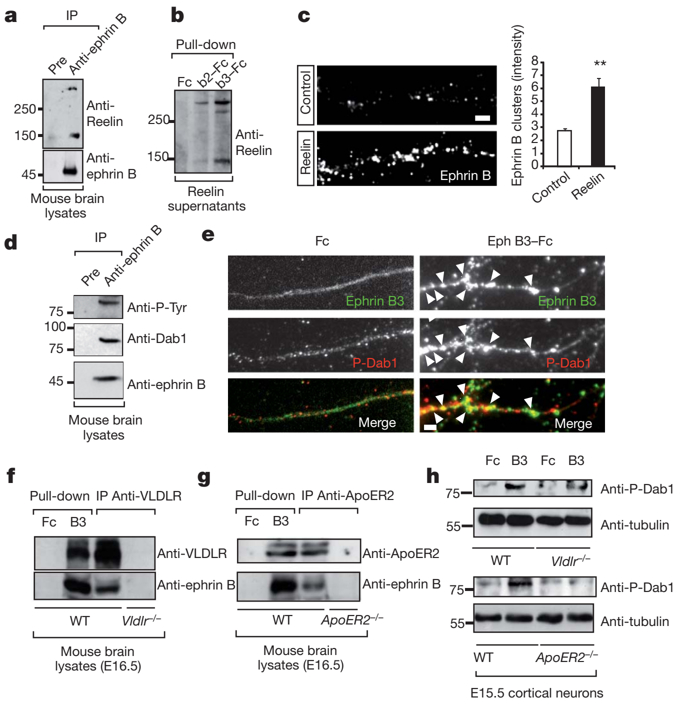


Figure 3 | Ephrin B2 and ephrin B3 ligands bind to Reelin and cluster VLDLR and ApoER2 signalling machinery. **a**, Reelin co-precipitates with ephrin B from embryonic mouse brain (E16.5) lysates. Pre, pre-immune serum. **b**, The extracellular domain of ephrin B2 and ephrin B3 binds to Reelin. Supernatants of 293 cells expressing Reelin were incubated with ephrin B2–Fc, ephrin B3–Fc or Fc as a control. **c**, Reelin clusters ephrin Bs in cortical neurons (E16.5 + 5 days in culture (DIC)). Quantification is based on signal intensity on dendritic branches (s.e.m., $n = 100$–150). **d**, Dab1 is in complex with ephrin B. Lysates from E16.5 brain were immunoprecipitated with anti-ephrin B antibody and analysed with anti-Dab1, anti-P-Tyr and anti-ephrin B antibodies. **e**, Stimulation of ephrin B3 with Eph B3–Fc induces clustering and phosphorylation of Dab1 in ephrin B3 patches in primary cortical neurons (E16.5 + 5 DIC) (arrowheads). **f**, **g**, Ephrin B3 co-immunoprecipitates with VLDLR and ApoER2. Lysates from E16.5 brains were used for Eph B3–Fc pull-downs and immunoprecipitation with anti-VLDLR or anti-ApoER2 antibodies. Fc, *Vldlr*$^{-/-}$ and *ApoER2*$^{-/-}$ brains were used as negative controls. **h**, *Vldlr*$^{-/-}$ and *ApoER2*$^{-/-}$ cortical neurons (E15.5 + 5 DIC) showed reduced P-Dab1 levels upon stimulation with Eph B3–Fc. The reduction in P-Dab1 was assessed by western blot analysis. Scale bars, 2 µm. $^{**}P < 0.01$.

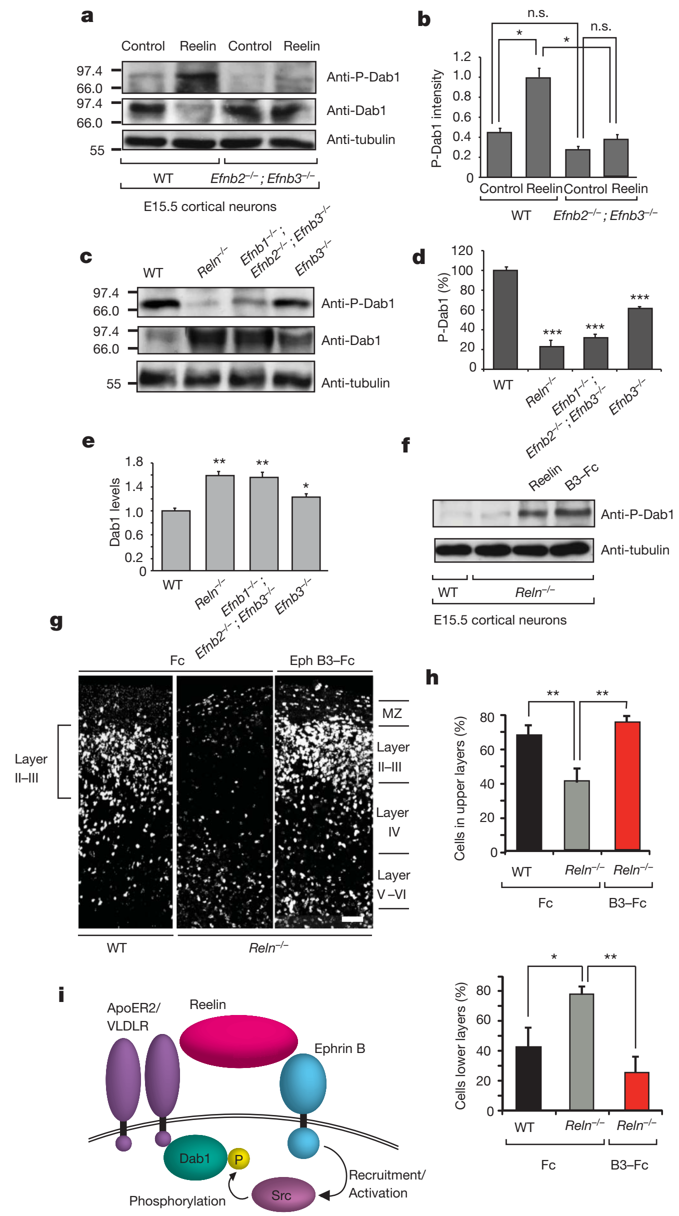


Figure 4 | Ephrin Bs are required for Reelin signalling *in vivo* and rescue the *reeler* defects. **a**, Ephrin Bs are required for Reelin-induced Dab1 phosphorylation. Mouse cortical neurons (E15.5) from double ephrin B2 and ephrin B3 knockouts are not able to phosphorylate Dab1 following stimulation with Reelin. **b**, Quantification of P-Dab1 signal intensity is shown. **c**, Dab1 phosphorylation is impaired *in vivo* in the *Efnb3*$^{-/-}$ and *Efnb1*$^{-/-}$; *Efnb2*$^{-/-}$; *Efnb3*$^{-/-}$ mice. **d, e**, Quantification of percentage of P-Dab1 signal (**d**) as well as Dab1 levels (**e**) is shown. **f**, Stimulation of *Reln*$^{-/-}$ cortical neurons (E15.5) with Eph B3–Fc is sufficient to induce Dab1 phosphorylation. Levels of Dab1 phosphorylation were assessed by western blot analysis. **g**, **h**, Activation of ephrin B3 in *Reln*$^{-/-}$ organotypic cortical cultures is sufficient to rescue the *reeler* phenotype. Pregnant females were injected with BrdU at E16.5 to label later born neurons that will constitute layers II–III. The correct positioning of BrdU$^+$ neurons was re-established following stimulation of slice cultures during 2 days with Eph B3–Fc. Distribution of BrdU$^+$ neurons is quantified both in the upper and lower cortical layers (s.e.m., $n = 24$–36) (**h**). **i**, Model for the function of ephrin Bs on Reelin signalling. Reelin binds to ephrin Bs and to the Reelin receptors. Ephrin Bs recruit and activate Src kinases which in turn phosphorylate the Dab1 associated to VLDLR and ApoER2 at the membrane. Scale bar, 50 µm. MZ, marginal zone; n.s., not significant. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

crucial components of the Reelin signalling machinery (Fig. 4i) that are required for the function of Reelin in neuronal migration. The regulation of Reelin signalling by ephrin Bs represents a further example of how ephrin Bs regulate the action of other membrane receptors. Ephrin Bs regulate cerebellar granule cell migration by controlling SDF-1-mediated activation of the G protein-coupled chemokine receptor CXCR4 (ref. 26). We have shown previously that ephrin B2 regulates the trafficking of AMPA receptors at the synapse, thereby controlling synaptic transmission[15]. In the vasculature our recent study has unravelled that ephrin B2 controls the internalization and activation of VEGFR2 during vascular development and tumour angiogenesis[13]. Now we provide evidence that ephrin Bs induce the formation of a macromolecular complex required for Src recruitment/activation and Reelin signalling. Taken together, ephrin Bs emerge now as general regulators of cellular signalling involved in the physical recruitment of signalling machineries to achieve proper spatial and temporal activation of other receptors implicated in a broad array of functions. Reelin signalling is required for proper nervous system function not only during development, but also during synaptic plasticity in the adult brain[3]. Importantly, we show that the activation of ephrin Bs is sufficient to rescue the absence of Reelin, indicating a major role of ephrin Bs regulating Reelin signalling and opening avenues to explore further potential therapeutic targets for neurological disorders associated with the loss of Reelin protein.

METHODS SUMMARY

Animals. Compound mice were generated as described in Methods. All the animal experiments were conducted under the institutional guidelines and were approved by the Hessen Animal Ethics Committees.

BrdU injection and immunohistochemistry. Pregnant mice were injected intraperitoneally with BrdU (Roche Diagnostics) in 0.9% NaCl (50 mg per kg) at E12.5, E15.5 and E17.5. Postnatal mice were perfused transcardially under anaesthesia with 4% paraformaldehyde (PFA) in PBS. Embryonic brains were immersion fixed in 4% PFA at 4 °C overnight. To characterize the migrating cells sections were treated for BrdU labelling as described previously[27]. Coronal sections of fixed brains were subjected to immunohistochemistry as described in Methods and images were acquired using a confocal microscope (Leica TCS SP2). Quantitative measurements were performed using MetaMorph software (Molecular Devices).

Stimulation of neurons. Cortical neurons from E14.5 were grown for 5 days *in vitro* and stimulated either with Eph B3–Fc chimaeras or with Reelin, using Fc or GFP as control conditions, respectively. Reelin and GFP were obtained from stable 293-HEK cell lines, as explained in Methods. Images were acquired using a digital camera (SpotRT; Diagnostic Instruments) attached to an epifluorescence microscope (Zeiss) equipped with a ×63 objective (Plan-Apochromat; Zeiss). All quantitative measurements were performed using MetaMorph (Molecular Devices).

Organotypic cultures. Cortical slices were prepared from embryonic mice (E14–E16) and cerebellar slices from P5 mice as described previously[28,29] and mounted onto porous nitrocellulose filters (0.45 μm; pore size, Millipore). Cortical and cerebellar sections were incubated in a humidified incubator supplied with 5% CO_2 at 35 °C. After 2–4 h recovery, stimulation and staining of cultured cortical slices were performed as described previously[30]. Cerebellar slices were electroporated (CUY21, NepaGene) with pGFP (Clontech), incubated 4 days and mounted for imaging.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 May 2010; accepted 25 January 2011.
Published online 3 April 2011.

1. D'Arcangelo, G. *et al.* Reelin is a ligand for lipoprotein receptors. *Neuron* **24,** 471–479 (1999).
2. Hong, S. E. *et al.* Autosomal recessive lissencephaly with cerebellar hypoplasia is associated with human *RELN* mutations. *Nature Genet.* **26,** 93–96 (2000).
3. Herz, J. & Chen, Y. Reelin, lipoprotein receptors and synaptic plasticity. *Nature Rev. Neurosci.* **7,** 850–859 (2006).
4. Rakic, P. Specification of cerebral cortical areas. *Science* **241,** 170–176 (1988).
5. Bystron, I., Blakemore, C. & Rakic, P. Development of the human cerebral cortex: Boulder Committee revisited. *Nature Rev. Neurosci.* **9,** 110–122 (2008).
6. Walsh, C. A. & Goffinet, A. M. Potential mechanisms of mutations that affect neuronal migration in man and mouse. *Curr. Opin. Genet. Dev.* **10,** 270–274 (2000).
7. Tissir, F. & Goffinet, A. M. Reelin and brain development. *Nature Rev. Neurosci.* **4,** 496–505 (2003).
8. Howell, B. W., Herrick, T. M. & Cooper, J. A. Reelin-induced tyrosine phosphorylation of Disabled 1 during neuronal positioning. *Genes Dev.* **13,** 643–648 (1999).
9. Bock, H. H. & Herz, J. Reelin activates SRC family tyrosine kinases in neurons. *Curr. Biol.* **13,** 18–26 (2003).
10. Arnaud, L., Ballif, B. A., Forster, E. & Cooper, J. A. Fyn tyrosine kinase is a critical regulator of disabled-1 during brain development. *Curr. Biol.* **13,** 9–17 (2003).
11. Costa, E. *et al. REELIN* and schizophrenia: a disease at the interface of the genome and the epigenome. *Mol. Interv.* **2,** 47–57 (2002).
12. Botella-Lopez, A. *et al.* Reelin expression and glycosylation patterns are altered in Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **103,** 5573–5578 (2006).
13. Sawamiphak, S. *et al.* Ephrin-B2 regulates VEGFR2 function in developmental and tumour angiogenesis. *Nature* **465,** 487–491 (2010).
14. Segura, I., Essmann, C. L., Weinges, S. & Acker-Palmer, A. Grb4 and GIT1 transduce ephrinB reverse signals modulating spine morphogenesis and synapse formation. *Nature Neurosci.* **10,** 301–310 (2007).
15. Essmann, C. L. *et al.* Serine phosphorylation of ephrinB2 regulates trafficking of synaptic AMPA receptors. *Nature Neurosci.* **11,** 1035–1043 (2008).
16. Palmer, A. *et al.* EphrinB phosphorylation and reverse signaling: regulation by Src kinases and PTP-BL phosphatase. *Mol. Cell* **9,** 725–737 (2002).
17. Caviness, V. S. Jr & Rakic, P. Mechanisms of cortical development: a view from mutations in mice. *Annu. Rev. Neurosci.* **1,** 297–326 (1978).
18. Caviness, V. S. Jr. Neocortical histogenesis in normal and reeler mice: a developmental study based upon [^{3}H]thymidine autoradiography. *Brain Res.* **256,** 293–302 (1982).
19. Hevner, R. F. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29,** 353–366 (2001).
20. Britanova, O. *et al.* Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. *Neuron* **57,** 378–392 (2008).
21. Sheppard, A. M. & Pearlman, A. L. Abnormal reorganization of preplate neurons and their associated extracellular matrix: an early manifestation of altered neocortical development in the reeler mutant mouse. *J. Comp. Neurol.* **378,** 173–179 (1997).
22. Forster, E. *et al.* Reelin, Disabled 1, and β_1 integrins are required for the formation of the radial glial scaffold in the hippocampus. *Proc. Natl Acad. Sci. USA* **99,** 13178–13183 (2002).
23. Liebl, D. J., Morris, C. J., Henkemeyer, M. & Parada, L. F. mRNA expression of ephrins and Eph receptor tyrosine kinases in the neonatal and adult mouse central nervous system. *J. Neurosci. Res.* **71,** 7–22 (2003).
24. Angrand, P. O. *et al.* Transgenic mouse proteomics identifies new 14-3-3-associated proteins involved in cytoskeletal rearrangements and cell signaling. *Mol. Cell. Proteomics* **5,** 2211–2227 (2006).
25. Trommsdorff, M. *et al.* Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97,** 689–701 (1999).
26. Lu, Q., Sun, E. E., Klein, R. S. & Flanagan, J. G. Ephrin-B reverse signaling is mediated by a novel PDZ-RGS protein and selectively inhibits G protein-coupled chemoattraction. *Cell* **105,** 69–79 (2001).
27. Hack, I. *et al.* Divergent roles of ApoER2 and Vldlr in the migration of cortical neurons. *Development* **134,** 3883–3891 (2007).
28. Alifragis, P., Parnavelas, J. G. & Nadarajah, B. A novel method of labeling and characterizing migrating neurons in the developing central nervous system. *Exp. Neurol.* **174,** 259–265 (2002).
29. Ghoumari, A. M., Wehrle, R., Bernard, O., Sotelo, C. & Dusart, I. Implication of Bcl-2 and Caspase-3 in age-related Purkinje cell death in murine organotypic culture: an *in vitro* model to study apoptosis. *Eur. J. Neurosci.* **12,** 2935–2949 (2000).
30. Nadarajah, B., Brunstrom, J. E., Grutzendler, J., Wong, R. O. & Pearlman, A. L. Two modes of radial migration in early development of the cerebral cortex. *Nature Neurosci.* **4,** 143–150 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank R. Klein for the ephrin B2$^{lox/lox}$ and ephrin B1$^{lox/lox}$ mice, N. Gale for the ephrin B3$^{-/-}$ mouse, J. Herz for the ApoER2$^{-/-}$ mouse, M. Goetz for 293-cells expressing Reelin or GFP, I. Segura for help with the ephrin B triple mutants, U. Bauer, O. Yildiz, C. Saygi, F. Kasikci, F. Voss and K. Happich for technical support, K. Hong and A. Griciuc for early contributions to this project, M. Frotscher for helpful discussions and I. Dikic and T. Acker for critically reading the manuscript. We acknowledge the Max Planck Institute of Neurobiology in Martinsried (Germany) for support and the use of equipment and animal facilities in the early stages of this project. This work was supported by grants from the Deutsche Forschungsgemeinschaft (AC180/2-1, 2-2 to A.A.-P.) and the Clusters of Excellence "Macromolecular Complexes (CEF)" (EXC115) and ECCPS (EXC147) at the University Frankfurt.

Author Contributions A.S. designed experiments and performed all the phenotypic characterization of the mouse mutants, the cortical neuron culture, organotypic slice culture rescue experiments and electroporation of cerebellar acute slice cultures. S.P. performed some of the neuronal culture experiments, organotypic slice culture rescue experiments and the biochemistry. A.W. and K.B. performed biochemistry for panels **d** and **a** in Fig. 4. A.A.-P. performed the initial biochemistry on the binding of ephrin B ligands to Reelin, designed experiments, interpreted results and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.A.-P. (Acker-Palmer@bio.uni-frankfurt.de).

METHODS

Animals and genotyping. Heterozygous Thy1–GFP, VLDLR and *reeler* mice were obtained from Jackson Laboratories and genotyped as suggested by the distributor. Single ephrin B3 knockout mice were provided by N. Gale (Regeneron Pharmaceuticals) and genotyped as described previously[31]. ApoER2 knockout mice were kindly provided by Joachim Herz and genotyped as described previously[32]. The generation of conditional ephrin $B2^{lox/lox}$ and the ephrin $B1^{lox/lox}$ knockout mice has been described[33,34]. To generate ephrin B3 or ephrin B2 compound mice, heterozygous *reeler* ($Reln^{+/-}$) mice were crossed with homozygous ephrin B3 knockout animals or $Nes\text{-}cre^{+}$; ephrin $B2^{lox/lox}$ and also with Thy1–GFP mice. Triple ephrin B1, B2, B3 knockout mice were generated by genetic crosses performed using single ephrin $B1^{lox/lox}$, ephrin $B2^{lox/lox}$ and ephrin B3 mouse lines. All the animal experiments were conducted under the institutional guidelines and were approved by the Hessen Animal Ethics Committees.

Immunohistochemistry (IHC) and tissue preparation. Postnatal mice were perfused transcardially under anaesthesia with 0.1 M phosphate buffered saline (PBS) containing 4% paraformaldehyde (PFA). Brains of young postnatal (P) mice (P0–P5) were immersion fixed in 4% PFA at 4 °C overnight. Coronal sections (50 µm) were cut using a vibratome (VT 1200S, Leica Instruments). Embryonic brains were fixed overnight in 4% PFA, cryoprotected in 30% sucrose in PBS, embedded in TissueTek OCT compound (Sakura Finetek Europa) and cut frozen coronally on a sliding microtome (Leica Instruments) at 20 µm thickness. For IHC, vibratome sections were washed in PBS, blocked in 1% BSA/0.5% Tween-20/PBS for 1 h at room temperature and incubated with primary antibodies (rabbit anti-calbindin D-28K (1:1,000), rabbit anti-Calretinin (1:1,000), rabbit anti-Dab1 (1:1,000), mouse anti-Reelin (1:1,000), mouse anti-NeuN (1:1,000), rabbit anti-Tbr1 (1:1,000) (all from Chemicon), rabbit anti-ApoER2 (1:500, Abcam), mouse anti-CSPG (1:400, Sigma), rabbit anti-GFAP (1:1,000, DAKO), goat anti-Brn1 (C-17) (1:100, Santa Cruz), rabbit anti-Dab1 (1:500, Chemicon), rabbit anti-Phospho-Dab1 (Tyr232) (1:1,000; Cell Signaling), rabbit anti-ApoER2 (1:1,000, Abcam), goat anti-VLDLR (1:200, R&D), human Eph B3–Fc, Eph B2–Fc chimaeras (1:100, R&D Systems), rat anti-BrdU (1:200, AbD Serotec), rabbit anti-SATB2 (1:500, Abcam)) for 24–48 h at 4 °C in the same solution. Sections were then washed and incubated for 2 h with either biotinylated secondary antibodies that were viewed by the ABCkit (Vector Laboratories) with diaminobenzidine (DAB) or with fluorescence-conjugated secondary antibodies (1:200, Jackson ImmunoResearch). Cryosections, collected on gelatin coated glass microscope slides, were first washed with PBS for 5 min, then boiled in 0.1% sodium citrate buffer (pH 8.0) 15 min in a microwave for antigen retrieval and placed in a humidified chamber for the rest of the staining procedure. Sections were blocked for 1 h at room temperature in 10% serum, 0.2% Triton X-100 PBS and incubated overnight with primary antibody at 4 °C in the same solution. After washing three times with PBS for 10 min, slices were incubated with fluorescent-labelled secondary antibody for 2 h at room temperature. All sections were coverslipped with Vectashield mounting medium with DAPI (Vector Laboratories).

BrdU injection and staining. For birth dating studies, the day of the vaginal plug was considered embryonic day (E) 0.5. The timed-pregnant female mice ($n = 2$–4 for each developmental point) were injected intraperitoneally at E12.5, E15.5 and E17.5 with a single pulse (50 mg per kg body weight) of 5-bromodeoxyuridine (BrdU) (5 mg ml^{-1} dissolved in 0.9% NaCl). The mice were perfused at P20 and BrdU staining was performed on 50-µm sections as described previously[27]. Briefly, sections were treated with 2 M HCl for 1 h, rinsed in 0.1 M sodium borate buffer (pH 8.4) to neutralize the residual acid for 10 min and after blocking 1 h in blocking solution (5% serum, 0.2% Triton X-100 PBS) at room temperature, incubated for two nights with monoclonal mouse anti-BrdU (1:500, Chemicon) at 4 °C. After washing the sections were incubated with Alexa-Fluor-488 secondary antibodies at room temperature for 2 h and embedded in fluorescent mounting medium (Dako).

Immunocytochemistry. Cortical and hippocampal neuron cultures were prepared from embryonic mouse brains (E15.5–E17.5) as described[15,35]. Briefly, cortices or hippocampi were removed and digested with 0.25% trypsin containing 1 mM EDTA (Invitrogen) for 20 min at 37 °C with gentle shaking. 30,000 to 70,000 (24-well plates) or 10^6 (6-cm plates) neurons were plated on coverslips or tissue culture plates, respectively, coated with poly-D-lysine (1 mg ml^{-1}) and laminin (5 µg ml^{-1}), in neurobasal medium supplemented with B27 containing 0.5 mM L-glutamine (all reagents from Invitrogen). Neurons were grown 5 days *in vitro* (DIV) and stimulated. Stimulation and staining of neuronal cortical and hippocampal cultures were performed as described[35]. For stimulation, unclustered recombinant mouse Eph B3–Fc, Eph B2–Fc chimaeras (R&D Systems) or Fc (Jackson ImmunoResearch) were pre-clustered for 1 h at room temperature using goat anti-human IgG (Jackson ImmunoResearch) as described previously[16]. For colocalization of P-Dab1, VLDLR and ApoER2 with ephrin clusters neurons were stimulated with Fc or Eph B3–Fc for 1 h. For inhibition of Src family kinases 10 µM SU6656 (Calbiochem) was added to the medium 1 h before the stimulation with Eph B3–Fc. The neurons were then fixed with 4% PFA and 4% sucrose in PBS for 12 min at 4 °C, rinsed twice with PBS, incubated with 50 mM NH_4Cl in PBS for 10 min at 4 °C and rinsed twice again before permeabilization for 5 min with ice cold 0.1% Triton X-100 in PBS. After washing with PBS, cells were blocked for 30 min at room temperature in blocking solution (2% bovine serum albumin, 4% serum (Jackson ImmunoResearch)) and incubated with primary (rabbit anti-ephrin B (C18) (1:100, Santa Cruz), rabbit anti-Phospho-Dab1 (Tyr 232) (1:1,000, Cell Signaling), goat anti-VLDLR (1:20, R&D), rabbit anti-ApoER2 (1:500, Sigma), goat anti-human Fc (1:50, Jackson Immuno Research)) and secondary antibodies for 60 and 30 min at room temperature, respectively. Samples were mounted using the Gel/Mount anti-fading medium (Biomeda).

Immunoblotting, immunoprecipitation and pull-down experiments. For immunoblotting, protein samples were separated by 7.5% SDS–PAGE and transferred to 0.45-µm nitrocellulose membranes (Schleicher & Schuell). For immunoprecipitation and direct western blot analysis cells or tissue from mouse brain were lysed in LBA lysis buffer (50 mM Tris HCl buffer, pH 7.5, 0.5% Triton X-100, 150 mM NaCl, 10 mM sodium pyrophosphate, 20 mM NaF, 1 mM sodium orthovanadate and 1% Complete (Roche)) and centrifuged at 10,000*g* for 10 min. Antibodies were pre-bound for 1 h to 10 µl protein A- or protein G-Sepharose beads (Pharmacia) and incubated with the lysates for 2 h at 4 °C and then applied to the gels. For western blot analysis, the following primary antibodies were used: rabbit anti-ephrin B (C-18) (1:200, Santa Cruz), mouse anti-phosphotyrosine 4G10 (1:1,000, Upstate), mouse anti-Reelin (1:1,000, Chemicon), rabbit anti-Dab1 (1:1,000, Chemicon), goat anti-Dab1 (1:200, Santa Cruz), rabbit anti-phospho-Dab1 (Tyr 232) (1:1,000; Cell Signaling), rabbit anti-ApoER2 (1:1,000, Abcam), goat anti-VLDLR (1:500, R&D), mouse anti-tubulin (1:50,000, Molecular Probes). The secondary antibodies used were horseradish peroxidase-conjugated goat anti-mouse, donkey anti-goat and goat anti-rabbit IgG (1:3,000, Jackson ImmunoResearch). Ephrin B2–Fc and ephrin B3–Fc pull-down experiments were performed as described[16]. Western blots are representative of at least three independent experiments. For stimulation, we used recombinant mouse Eph B3–Fc (R&D Systems), human Fc (Jackson ImmunoResearch) and Reelin and GFP (as a control) from supernatants from 293 cells expressing Reelin and GFP respectively.

Preparation of Reelin-containing and control supernatants. To obtain Reelin-enriched supernatants and control supernatants, the incubation medium (DMEM, 10% fetal calf serum, 0.360 g l^{-1} G418, 2 mM L-glutamine (all from Invitrogen), 0.1 g l^{-1} penicillin/streptomycin (PAA)) from Reelin-transfected 293-HEK cells or GFP-transfected control 293-HEK cells (a gift from M. Goetz) was replaced by serum-free medium (Optimem-1 with GlutaMAX (Gibco), 0.1 g l^{-1} penicillin/streptomycin (PAA), 2 mM L-glutamine (Invitrogen)) and cells were incubated for 2 days at 37 °C, 5% CO_2. The conditioned medium was collected and concentrated ~ 20-fold by centrifugation using Amicon Ultra filters (100K, Millipore) and the Reelin content (absence of Reelin in control cell supernatants) was confirmed by western blotting using mouse anti-Reelin antibody (1:1,000, Chemicon).

Organotypic cultures. Cerebellar slices from P6 mice and brain slices from embryonic mice (E14–E16) were prepared as described previously[28,29]. Briefly, brains were dissected out and embedded in 3% low melting point agarose (Sigma) and sectioned immediately in ice-cold preparation medium (MEM (Invitrogen), 0.1 g l^{-1} penicillin/streptomycin (PAA), 25 mM HEPES (Invitrogen), 2 mM L-glutamine (Invitrogen), 0.45% glucose and 1 N NaOH, pH 7.4), at 300 µm using a vibratome (VT 1200S, Leica Instruments). Coronal slices of embryonic brains and cerebellar sections were mounted onto porous nitrocellulose filters (0.45 µm; Millipore) and transferred to 6-well culture plates. Slices were allowed to recover for 1 h in defined culture medium with DMEM (Invitrogen), 5% heat-inactivated horse serum (Sigma), 1× N-2 (Gibco), 100 µM L-glutamine (Invitrogen), 2.4 g l^{-1} glucose, and 0.1 g l^{-1} penicillin/streptomycin (PAA) 1× B27 (Gibco), 25% BME (Gibco) and 25 mM HEPES, with 5% CO_2 and 35 °C. The amount of medium added to each well was that required to cover the slices with a meniscus of fluid. Organotypic cultures of embryonic brains were stimulated with Eph B3–Fc or with Fc, as control, at 35 °C in a humidified incubator with 5% CO_2 for 2–4 days. We performed electroporation on cerebellar organotypic cultures. Green fluorescent protein (pGFP; Clontech) expression vector was dissolved in PBS at a concentration of 0.5 µg $µl^{-1}$. Immediately before use, Fast Blue solution (0.1% in PBS, Sigma) was added to the plasmid solution at a ratio of 1:10 to monitor the injection. Plasmid solution (~1 µl) was applied onto the cerebellar cultures using a glass micropipette. The electronic pulses (20 V, 50 ms, five times, 950 ms intervals) were then delivered using a sequence wave electroporator (CUY21, Nepa Gene). Cultured cortical and cerebellar slices were fixed with 4% PFA in 0.1 M PBS and 5% sucrose, washed several times with TBS and processed for IHC as described previously[30]. To characterize the phenotype of migrating cells, sections were treated for BrdU labelling as described previously[27].

Imaging and quantifications. Images of immunocytochemical staining were acquired using a digital camera (SpotRT; Diagnostic Instruments) attached to an epifluorescence microscope (Zeiss) equipped with a ×63 objective (Plan-Apochromat; Zeiss). Immunohistochemical staining done on vibratome sections were acquired using confocal microscope (Leica TCS SP2) equipped with ×20 and ×40 objectives. Immunohistochemical staining performed using ABCkit were acquired using a stereomicroscope (Leica M165 FC). All quantitative measurements were performed using MetaMorph software (Molecular Devices). In the cerebellum the branching defects of the Purkinje cells were analysed by setting a radial line from the Purkinje cell soma to the granular cell layer in the cerebellum. The distance was measured from the first branching point that occurs in the dendritic tree of every secondary branch of a Purkinje cell to the radial line. BrdU experiments were performed on a total number of 34 embryonic brains coming from injected females at time points E12.5, E15.5 and E17.5. From the different genotypes between 23 and 58 sections were counted with a total number of neurons that ranged between 448 and 1,022 per genotype. The images were taken consistently from all the genotypes from the S1 (primary somatosensory cortex) and the PPtA (posterior parietal association area). Quantification of Dab1 phosphorylation was based on fluorescence intensities. The intensity of P-Dab1 was calculated for dendrite stretches of 100–200 μm imaged on at least 10 different treated neurons ($n = 50$–100). All the analysis of cortex, hippocampus and cerebellum was performed in 3–4 animals per group. Quantification of cell densities at various areas of cortex was calculated for 100–200 μm^2 cortical area in 10–30 sections per animal. Sholl analysis was performed to characterize the morphological characteristics of neurons by counting the number of dendrite intersections for concentric circles starting at the centre of the cell body, from 20–30 neurons from each pyramidal layer of 3–4 animals. Rescue experiments in organotypic slice cultures were performed with four slices per condition in three different experiments. Quantification of western blots was performed by measuring band sizes and intensities using Photoshop (Adobe). Statistical analysis was performed using Microsoft Excel and student's *t*-tests were used to assess statistical significance of the differences between measurements.

31. Kullander, K. *et al.* Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. *Genes Dev.* **15,** 877–888 (2001).
32. Trommsdorff, M. *et al.* Reeler/Disabled-like disruption of neuronal migration in knockout mice lacking the VLDL receptor and ApoE receptor 2. *Cell* **97,** 689–701 (1999).
33. Grunwald, I. C. *et al.* Hippocampal plasticity requires postsynaptic ephrinBs. *Nature Neurosci.* **7,** 33–40 (2003).
34. Compagni, A., Logan, M., Klein, R. & Adams, R. H. Control of skeletal patterning by ephrinB1-EphB interactions. *Dev. Cell* **5,** 217–230 (2003).
35. Zhang, Y. & Bhavnani, B. R. Glutamate-induced apoptosis in primary cortical neurons is inhibited by equine estrogens via down-regulation of caspase-3 and prevention of mitochondrial cytochrome c release. *BMC Neurosci.* **6,** 13 (2005).

LETTER

doi:10.1038/nature09976

TRIM5 is an innate immune sensor for the retrovirus capsid lattice

Thomas Pertel[1], Stéphane Hausmann[1], Damien Morger[2], Sara Züger[2], Jessica Guerra[1], Josefina Lascano[1], Christian Reinhard[1], Federico A. Santoni[1], Pradeep D. Uchil[3], Laurence Chatel[4], Aurélie Bisiaux[5], Matthew L. Albert[5], Caterina Strambio-De-Castillia[1], Walther Mothes[3], Massimo Pizzato[1], Markus G. Grütter[2] & Jeremy Luban[1]

TRIM5 is a RING domain-E3 ubiquitin ligase that restricts infection by human immunodeficiency virus (HIV)-1 and other retroviruses immediately following virus invasion of the target cell cytoplasm[1,2]. Antiviral potency correlates with TRIM5 avidity for the retrovirion capsid lattice[3,4] and several reports indicate that TRIM5 has a role in signal transduction[5–7], but the precise mechanism of restriction is unknown[8]. Here we demonstrate that TRIM5 promotes innate immune signalling and that this activity is amplified by retroviral infection and interaction with the capsid lattice. Acting with the heterodimeric, ubiquitin-conjugating enzyme UBC13–UEV1A (also known as UBE2N–UBE2V1), TRIM5 catalyses the synthesis of unattached K63-linked ubiquitin chains that activate the TAK1 (also known as MAP3K7) kinase complex and stimulate AP-1 and NFκB signalling. Interaction with the HIV-1 capsid lattice greatly enhances the UBC13–UEV1A-dependent E3 activity of TRIM5 and challenge with retroviruses induces the transcription of AP-1 and NF-κB-dependent factors with a magnitude that tracks with TRIM5 avidity for the invading capsid. Finally, TAK1 and UBC13–UEV1A contribute to capsid-specific restriction by TRIM5. Thus, the retroviral restriction factor TRIM5 has two additional activities that are linked to restriction: it constitutively promotes innate immune signalling and it acts as a pattern recognition receptor specific for the retrovirus capsid lattice.

To determine if TRIM5 contributes to signal transduction, the effect of ectopic human TRIM5α expression on transcriptional reporters in HEK-293 cells was examined. TRIM5 stimulated either of two luciferase reporters for AP-1 with a magnitude comparable to that of MAVS or the AP-1 transcription factor c-Jun (Fig. 1a and Supplementary Fig. 1a). TRIM5 also stimulated NF-κB (Fig. 1b) but minimally activated *IFNB1*-, or IRF3-dependent, luciferase reporters (Fig. 1c and Supplementary Fig. 1b and c). The TRIM5–cyclophilin A fusion protein from owl monkey[1] activated AP-1 and NF-κB to similar levels as human TRIM5α (Supplementary Fig. 1d, e). Although TRIM5 was not sufficient to activate *IFNB1*, induction of *IFNB1* by IRF3 was greatly enhanced by TRIM5 (Fig. 1c), consistent with the fact that *IFNB1* transcription requires NF-κB and AP-1, as well as IRF3 (Supplementary Fig. 1f)[9].

To determine if endogenous TRIM5 regulates AP-1 and NF-κB signalling pathways, the effect of *TRIM5* knockdown was assessed in myeloid cells. THP-1 cells were transduced with lentiviral vectors engineered to confer puromycin-resistance and to express RNA polymerase II (Pol II)-driven, microRNA-based short hairpin RNAs (shRNAs) targeting either *TRIM5* or control RNAs (Supplementary Fig. 2a–c). Pools of puromycin-resistant cells were generated with each knockdown vector and global expression profiles were assessed. The effect of *TRIM5* knockdown was extraordinarily specific in that, of 25,000 genes probed, only 33 were significantly decreased (Fig. 1d). The majority of these were NF-κB- and AP-1-responsive inflammatory mediators, 70% being inflammatory chemokines and cytokines (Supplementary Table 1).

Lipopolysaccharide (LPS), a pathogen-associated molecular pattern (PAMP) recognized by the pattern recognition receptor (PRR) TLR4-MD-2, activates AP-1 and NF-κB-signalling and this culminates in the expression of inflammatory genes like those perturbed by *TRIM5* knockdown[10,11]. Monocyte-derived dendritic cells (MDDC), macrophages (MDM) and THP-1 cells were challenged with LPS and induction of the AP-1- and NF-κB-dependent genes *CXCL9*, *CXCL10*, *CCL8*, *IL6*, *IL8* and *PTGS2* (also known as *COX2*), was found to be attenuated by *TRIM5* knockdown (Fig. 1e and f and Supplementary Fig. 2d and e). These results demonstrate that TRIM5 activates MAPK- and NF-κB-dependent genes and makes a major contribution to LPS signalling and gene induction (Supplementary Fig. 1f).

Given the contribution of TRIM5 to the production of inflammatory mediators by LPS, the effect of TRIM5 on the previously reported anti-HIV-1 activity of LPS[12] was examined. Transduction of MDDC, MDM or THP-1 macrophages by vesicular stomatitis virus (VSV) G-pseudotyped HIV-1 was blocked by LPS, by other PAMPs, and by type 1 IFN (Supplementary Fig. 3a–c). *TRIM5* mRNA increased tenfold in response to these factors (Supplementary Fig. 3d, e), but this increase was not sufficient for the anti-HIV-1 state (Supplementary Fig. 3f, g). Nonetheless, *TRIM5* knockdown rescued HIV-1 from LPS, although not from type 1 IFN, and the magnitude rescue correlated with the efficiency of *TRIM5* knockdown (Fig. 1g and Supplementary Fig. 4a and b). These phenotypes were indistinguishable from those observed with knockdown of *IRF3*, a critical transcription factor that acts proximal to *IFNB1* (ref. 10; Fig. 1h and Supplementary Fig. 4c). In contrast, knockdown of *STAT2*, a factor that acts downstream of the type I IFN receptor, blocked the anti-HIV-1 activity of either LPS or type 1 IFN (Fig. 1i and Supplementary Fig. 4d).

Rescue from LPS seems to be independent of capsid-recognition by TRIM5 in that *TRIM5* knockdown rescued a molecular clone of simian immunodeficiency virus (SIV_{MAC}), a retrovirus that differs greatly from HIV-1 in terms of its sensitivity to TRIM5-mediated restriction[1,2], as well as two non-retroviruses, the rhabdovirus vesicular stomatitis virus and the paramyxovirus Newcastle disease virus (Fig. 1j and Supplementary Fig. 4e–i). Although TRIM5 is not sufficient to activate IFNB1 (Fig. 1c), it promotes the first wave of innate immune signalling upstream of IFNB1 and thereby contributes to the antiviral state established by LPS (Supplementary Fig. 1f).

To understand how TRIM5 activates AP-1 and NF-κB, 20 candidate proteins, selected on the basis of signalling activity above the MAPK/NF-κB bifurcation in the LPS signalling pathway, were tested for the ability to immunoprecipitate with TRIM5. Strong signal was observed with TAK1, TAB2, and TAB3 (Fig. 2a and Supplementary Fig. 5a), all components of the TAK1 kinase complex that phosphorylates proximal MAPK and NF-κB kinases in response to LPS[11]. Like TRIM5, TAK1 potently activated AP-1 and modestly activated NF-κB (Fig. 2b). 5Z-7-oxozeaenol, a TAK1-inhibitor, blocked AP-1 induction by TRIM5 or

[1]Department of Microbiology and Molecular Medicine, University of Geneva, Geneva CH-1211, Switzerland. [2]Department of Biochemistry, University of Zurich, Zurich CH-8057, Switzerland. [3]Section of Microbial Pathogenesis, Yale University School of Medicine, New Haven, Connecticut 06536, USA. [4]Novimmune SA, Geneva CH-1228, Switzerland. [5]Institut Pasteur, Inserm U818, Paris 75724, France.

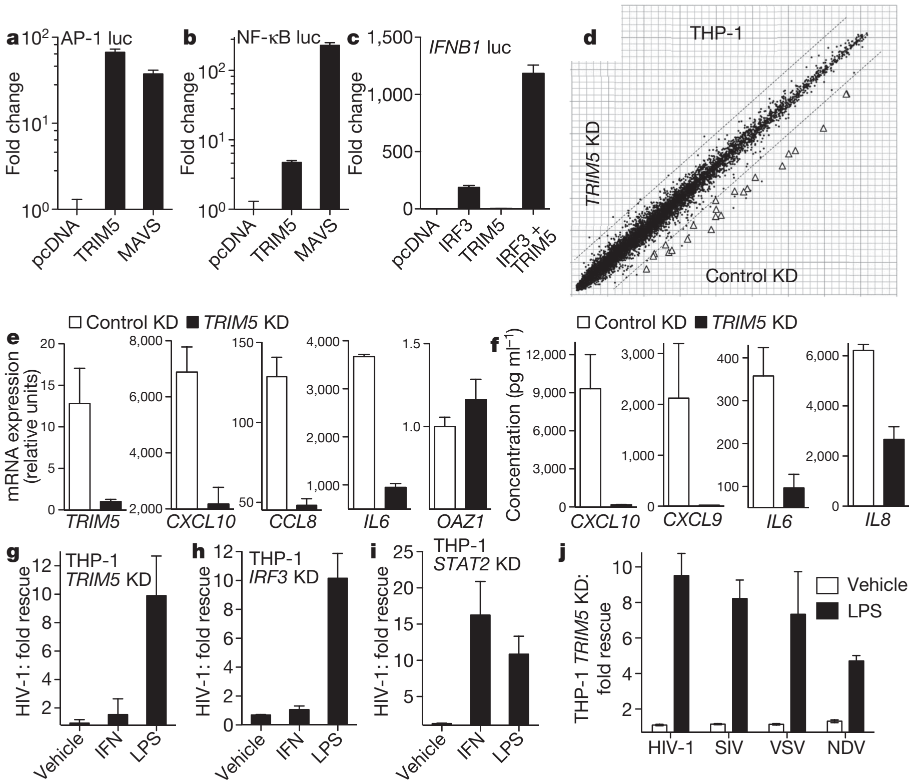


Figure 1 | TRIM5 promotes innate immune signalling. **a**–**c**, HEK-293 cells transfected with the indicated pcDNA-based expression plasmids and luciferase reporters for AP-1 (**a**), NF-κB (**b**) or *IFNB1* (**c**). Bars show mean luciferase activity ± s.d. ($n = 6$). **d**, Global expression profile comparing *TRIM5* knockdown (KD) to control KD THP-1 macrophages. Triangles indicate inflammatory genes significantly downregulated in *TRIM5* KD. **e**, qRT–PCR for the indicated mRNAs collected from MDDCs 2 to 8 h after LPS treatment, depending on the peak values for that gene. Shown are the means ± s.e.m ($n = 3$) relative to untreated cells. **f**, Concentration of the indicated proteins in the culture supernatant, 24 h after LPS treatment (mean ± s.d., $n = 3$). RNA and protein data are representative of at least three separate donors. **g**–**j**, THP-1 macrophages transduced with miR30-based lentivirus KD vectors targeting either *TRIM5* (**g** and **j**), *IRF3* (**h**), or *STAT2* (**i**), were treated for 24 h with the indicated compounds and challenged with VSV-G pseudotyped HIV-1 luciferase reporter virus (**g**–**i**) or with the indicated green fluorescent protein (GFP) reporter viruses (**j**). Data are expressed as fold-change compared to control KD cells, with s.e.m ($n = 4$). All data are representative of at least three independent experiments.

TAK-1 without effect on AP-1 induction by the downstream effector c-Jun (Fig. 2b). *TAK1* knockdown blocked AP-1 activation by TRIM5 (Fig. 2c), but not by c-Jun (Fig. 2c and Supplementary Fig. 5c). *TRIM5* knockdown blocked LPS-induced TAK1 autophosphorylation on threonine 187 (Fig. 2d), a post-translational modification required for TAK1 activation[11]. Like *TRIM5* knockdown, *TAK1* knockdown rescued HIV-1 from the LPS-induced antiviral state (Fig. 2e and Supplementary Fig. 5d), and either TAB2 or TAB3 acted synergistically with TRIM5 to activate AP-1 (Fig. 2f). These results indicate that TRIM5 and the TAK1 kinase complex cooperate to promote signal transduction, and given that TAK1 phosphorylates both IκB kinases (IKKs) and mitogen-activated protein kinase kinases (MKKs)[11], explains how TRIM5 activates both MAPK and NF-κB signalling pathways.

The well-characterized restriction of HIV-1 by owl monkey TRIMCyp[1,13] (a TRIM5–CypA fusion protein) was exploited to determine if TAK1 contributes to TRIM5-mediated, capsid-specific restriction. Pools of THP-1 cells were selected for puromycin-resistance after transduction with a bicistronic lentiviral vector encoding owl monkey TRIM5Cyp[13]. As shown previously, these cells were resistant to infection with wild-type HIV-1, but not to the HIV-1 G89V capsid mutant, and the infectivity of wild-type HIV-1 was rescued by cyclosporine[13] (Supplementary Fig. 5e). Control cells transduced with a vector bearing TRIM5Cyp(H436Q), a mutant that does not bind HIV-1 capsid and does not restrict HIV-1 (ref. 13), were infected with efficiency equal to that of cells transduced with the empty vector. THP-1 cells transduced with either wild-type or H436Q mutant TRIM5Cyp were then subjected to a second round of selection after transduction with miR30-based knockdown vectors targeting *TAK1* or luciferase control and expressing hygromycin-resistance. The pools of puromycin/hygromycin double-resistant THP-1 cells were then challenged with HIV-1. *TAK1* knockdown rescued HIV-1 transduction and nascent HIV-1 cDNA synthesis (Fig. 2g, h). This effect was specific to the cells with TRIM5Cyp-mediated restriction activity because *TAK1* knockdown had no effect on HIV-1 transduction in the non-restrictive, H436Q control cells (Fig. 2g).

The contribution of TAK1 to restriction of N-tropic murine leukemia virus (MLV) by human TRIM5α was examined using miR30-based knockdown vectors in THP-1, HeLa and HT1080 cells. Inhibition of both N-tropic and B-tropic MLV infection by the *TAK1* knockdown was observed, perhaps because, unlike HIV-1, infection with MLV is cell-cycle dependent[14], and these viruses were sensitive to growth inhibitory effects of the knockdown. This precluded assessment of capsid-specific effects on reporter gene transduction, although nascent viral cDNA synthesized after infection of THP-1 cells was rescued by the *TAK1* knockdown in an N-tropic MLV-specific manner (Supplementary Fig. 5g). Similar non-specific effects on MLV were observed after transfection of double stranded RNA (dsRNA) oligonucleotides targeting *TAK1*. Like HIV-1, equine infectious anaemia virus (EIAV) is a lentivirus that infects non-dividing cells, but it differs from HIV-1 in that it is relatively sensitive to human TRIM5α-mediated restriction[15]. Transfection of dsRNAs targeting *TAK1* rescued EIAV transduction almost to the same level as the *TRIM5* knockdown (Fig. 2i). These results indicate that TAK1 contributes to capsid-specific restriction mediated by TRIM5.

AP-1 induction by TRIM5 was impaired by mutants of the RING E3 ubiquitin (Ub)-ligase domain (Fig. 3a). This raised the question which

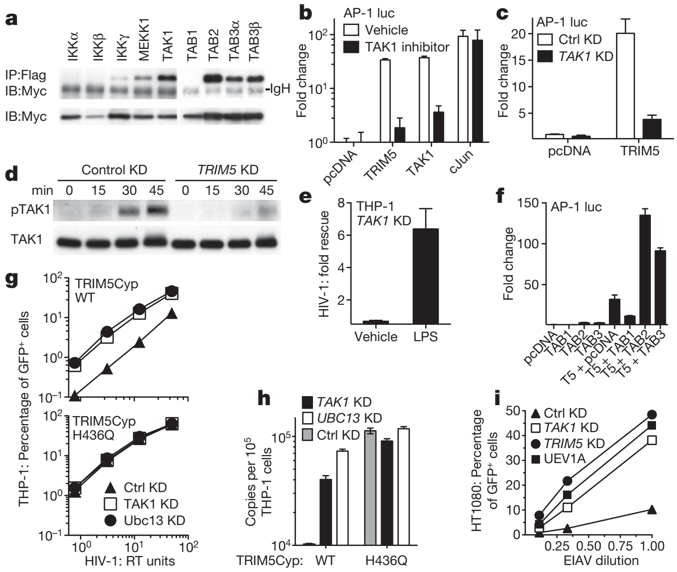


Figure 2 | The TAK1 kinase complex interacts biochemically and functionally with TRIM5. **a**, HEK-293T cells were co-transfected with Myc-tagged human TRIM5α and the indicated Flag-tagged constructs. Shown are immunoblots (IB) with anti-Myc antibody after immunoprecipitation with anti-Flag (upper panel), or of total cell lysate (bottom panel). **b**, **c**, and **f**, HEK-293 cells were transfected with the indicated pcDNA-based expression plasmids and an AP-1 luciferase reporter and show the effect of TAK1 inhibitor 5Z-7-oxozeaenol (**b**) or *TAK1* KD (**c**). *TAK1* KD and control KD THP-1 macrophages were treated with LPS for the indicated times and immunoblotted with anti-TAK1 antibody (lower panel) or anti-phospho-TAK1 antibody (upper panel) (**d**), or, cells were treated 24 h with LPS or vehicle and challenged with an HIV-1 luciferase reporter virus (**e**). The results in (**e**) are reported as fold rescue due to *TAK1* KD, with respect to control KD. **g** and **h**, THP-1 cells were transduced with lentiviral vectors encoding owl monkey TRIM5Cyp, either wild-type (WT) or the H436Q mutant. Pools of each were then transduced with lentiviral KD vectors targeting either *TAK1*, *UBC13* or control, and challenged with an HIV-1-GFP reporter vector. Infectivity was monitored by FACS (**g**) or by PCR for synthesis of full-length viral cDNA (**h**). **i**, HT1080 cells were transfected with dsRNA oligonucleotides targeting *TRIM5*, *TAK1* or *UEV1A* and challenged with EIAV-GFP reporter vector.

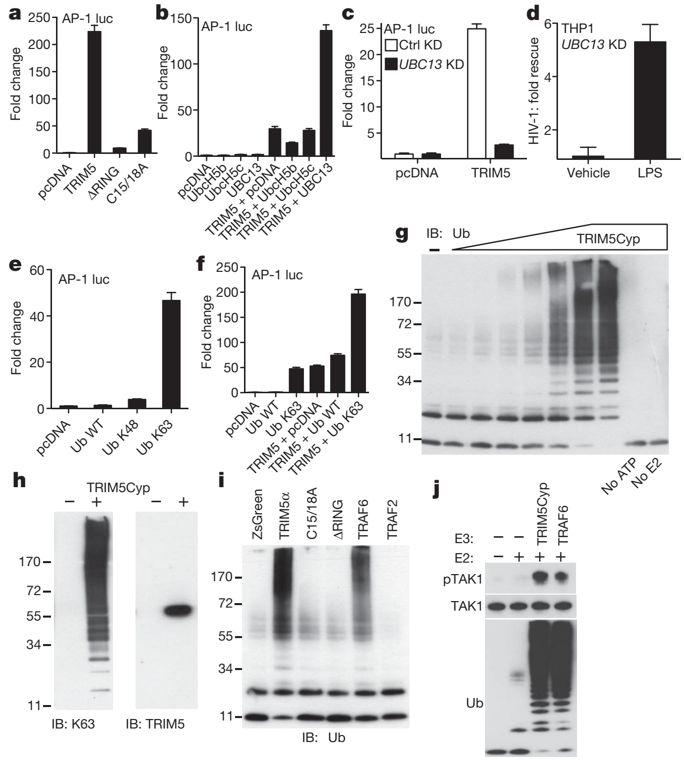


Figure 3 | TRIM5 acts with UBC13–UEV1A to synthesize free K63-linked Ub chains that activate TAK1. **a**–**c**, **e** and **f**, HEK-293 cells were transfected with an AP-1 luciferase reporter and the indicated pcDNA-based expression plasmids. Bars show mean ± s.d. ($n = 6$). In **c**, HEK-293 cells had stable *UBC13* KD or control KD. **d**, *UBC13* KD or control KD THP-1 macrophages were treated for 24 h with LPS or vehicle and challenged with an HIV-1 luciferase reporter virus. Shown is the fold rescue due to *UBC13* KD, with respect to the control KD. **g**-**j**, Products of *in vitro* reactions with ATP, Ub, UBE1, UBC13–UEV1A, and the indicated E3 Ub ligases were revealed by immunoblot for total Ub (**g**, **i**, and **j**), K63-linked Ub chains (left panel of **h**), or TRIM5 (right panel of **h**). E3 ubiquitin ligases included purified owl monkey TRIM5Cyp (**g**, **h**, and **j**), or the indicated Flag-tagged proteins immunoprecipitated from HEK-293T cells (**i**). **j**, *In vitro* Ub reactions like those in (**i**) were incubated with purified TAK1 kinase complex. Products were probed in immunoblot with the indicated antibodies.

of the many E2 Ub-conjugating enzymes might be relevant for TRIM5-mediated effects on signal transduction. Among candidate E2s, UBC13 synergized with TRIM5 to activate AP-1 (Fig. 3b). Interestingly, the TAK1 kinase complex is activated by the heterodimeric E2 UBC13–UEV1A[11]. Knockdown of *UBC13* or *UEV1A* severely blocked AP-1 activation by TRIM5 (Fig. 3c and Supplementary Fig. 6a–c), rescued HIV-1 from the LPS-induced antiviral state in THP-1 macrophages (Fig. 3d and Supplementary Fig. 6d), and rescued HIV-1 and EIAV from TRIM5-mediated restriction (Fig. 2g–i).

The UBC13–UEV1A E2 heterodimer is notable in that it generates K63-linked Ub chains that are unlinked to substrates; these free Ub chains multimerize and activate the TAK1 kinase complex via the Ub binding components, TAB2 and TAB3 (ref. 11). Ub in which all lysines except K63 are mutated to arginine (Ub K63) activated AP-1 and NF-κB (Fig. 3e and Supplementary Fig. 6e, f), and enhanced the ability of TRIM5 to activate AP-1 (Fig. 3f). K48-only Ub did not have these activities (Fig. 3e, f), nor did wild-type Ub, perhaps because of the dominance of competing Ub metabolic pathways and the tight regulation of K63 chains within cells[16]. These experiments indicate that the heterodimeric E2 UBC13–UEV1A and the K63-linked Ub chains that it produces have a role in TRIM5-mediated signalling.

Because TRIM5 interacted biochemically and functionally with TAK1, TAB2, TAB3, UBC13, UEV1A and K63-Ub, the ability of TRIM5 to synthesize K63-linked Ub chains was assessed. A purification protocol was established that yielded 0.5 mg of soluble, full-length, owl monkey TRIM5Cyp from 1 l of Sf9 cell culture (Supplementary Fig. 7). No procedure has been reported to date for the production of purified, full-length, recombinant TRIM5 protein[17]. Purified TRIM5Cyp was incubated with purified Ub, E1 and the E2 Ub-conjugases UBC13 and UEV1A, and reaction products were separated by SDS–polyacrylamide gel electrophoresis (SDS–PAGE). With increasing TRIM5Cyp concentration, monomeric Ub was progressively depleted and the yield of Ub chains increased (Fig. 3g and Supplementary Fig. 8a). Synthesis of Ub chains was ATP-dependent and required both UBC13 and UEV1A.

The Ub chains generated by TRIM5Cyp were detected with antibody specific for K63-linked Ub and immunoblot showed TRIM5Cyp to be a monomer with no detectable incorporation into the Ub chains (Fig. 3h). To obtain an independent assessment of their identity, reaction products were isolated by PAGE and analysed by matrix-assisted laser desorption/ionization and tandem mass spectrometry (Supplementary Figs 8a and 9a–c). These methods identified peptides corresponding to K63-linked Ub and failed to detect conjugates with other Ub lysines or peptides corresponding to TRIM5Cyp, confirming that reaction products were free, unattached K63 Ub chains. Additionally, synthesis of Ub chains was undetectable with a Ub mutant in which K63 was mutated to arginine (Supplementary Fig. 8b). Conversely, Ub was efficiently incorporated into chains when all lysines except K63 were mutated to arginine (Supplementary Fig. 8b), indicating that K63 was necessary and sufficient to form the Ub chains.

Human TRIM5α, produced by transfection of 293T cells and enriched by immunoprecipitation, catalysed the synthesis of free K63 Ub chains like those of TRIM5Cyp, in a RING domain-dependent manner (Fig. 3i and Supplementary Fig. 8c, d). It had at least as much activity as TRAF6 (Fig. 3i), an E3 Ub ligase previously reported to synthesize unattached K63 chains that activate TAK1 (ref. 11). TRAF2, a close paralogue of TRAF6 that does not interact with UBC13 (ref. 18), lacked activity (Fig. 3i).

Free K63-linked Ub chains generated by TRAF6 result in TAK1 autophosphorylation on threonine 187 (ref. 11), a modification required for TAK1 activation. To test the effect of K63-linked Ub chains generated by TRIM5 on TAK1 activation, the essential components of a TAK1 kinase complex, TAK1, TAB1 and TAB2 (ref. 11), were purified and combined (Supplementary Fig. 10). This complex was then incubated with Ub, UBC13–UEV1A, and either TRAF6 or purified owl monkey TRIM5Cyp. TAK1 phosphorylation was observed in response to the K63-linked Ub chains synthesized by either TRAF6 or by TRIM5Cyp (Fig. 3j). Kinase activity required the TAK1-associated TAB1, the Ub receptor TAB2, and UBC13–UEV1A (Fig. 3j). These experiments show that, like TRAF6 (ref. 11), TRIM5 synthesizes free K63-linked Ub chains that activate TAK1 autophosphorylation.

If TRIM5 were a PRR specific for the retroviral capsid lattice, infection with retroviruses would activate signalling, the magnitude of which would correlate with TRIM5 avidity for the capsid of the challenge virus. To determine if this is the case, myeloid cells were challenged with pairs of retroviruses that differ with respect to TRIM5 avidity for the capsid[3,4] and the subsequent induction of NF-κB- and MAPK-dependent genes was assessed. VSV G-pseudotyped N-tropic and B-tropic MLV vectors, normalized for exogenous reverse transcriptase activity and for titre on non-restrictive MDTF cells[19], were used to challenge THP-1 macrophages. The multiplicity of infection of the non-restricted B-tropic MLV on cycling THP-1 cells was 0.1. mRNA was harvested from the THP-1 cells and processed by reverse transcription and quantitative PCR (qRT–PCR). Greater induction of *PTGS2*, *CXCL10*, *CCL8* and *IL6* mRNA was observed after challenge with N-MLV than with B-MLV (Fig. 4a, b), in correlation with the higher avidity of human TRIM5α for the capsid of N-tropic MLV than for the capsid of B-tropic MLV[3]. *TRIM5* knockdown suppressed the higher inflammatory gene induction by N-MLV, indicating its dependence upon endogenous TRIM5 (Fig. 4b). Similar differential induction of *PTGS2*, *CXCL10*, *CCL8* and *IL6* mRNAs by N-tropic and B-tropic MLV was observed after challenge of MDDCs or MDMs (Supplementary Fig. 11).

Retroviral cDNA activates innate immune signalling under some conditions[20]. Restriction by human TRIM5α results in N-MLV cDNA levels that are an order of magnitude lower than for B-MLV cDNA[21] so the experiments described above might underestimate the effect of N-MLV capsid on TRIM5-mediated signalling. Therefore, MDDCs were challenged with matched pairs of virus-like particles (VLPs) devoid of the viral genome that serves as the reverse transcription template. VLPs bearing N-MLV capsid activated *CXCL9*, *CXCL10*, *IFIT1* and *IFIT2* mRNAs from 5- to 55-fold over the levels in untreated MDDCs (Fig. 4c). Inflammatory gene induction was not detected with VLPs bearing the unrestricted NB-MLV capsid[22] (Fig. 4c). As with the mRNA, soluble IL8, CCL5, CXCL9 and CXCL10 protein was differentially induced by N-MLV (Fig. 4d).

To determine if differential gene induction after retrovirus challenge was peculiar to N-tropic MLV, a similar experiment was performed with OMK, a kidney cell line from the owl monkey, *Aotus trivirgatus*. The TRIM5 orthologue in this species restricts HIV-1 but not SIV[1]. VSV G-pseudotyped HIV-1 and SIV vectors, normalized for exogenous reverse transcriptase activity and for titre on HeLa cells, were used to challenge OMK cells. The multiplicity of infection of the unrestricted SIV on OMK cells was 0.3. Among the MAPK- and NF-κB-dependent gene products that were detectable in this species using human probes, transcriptional activation of *PTGS2*, *IFIT1* and *IFIT2* mRNAs, and

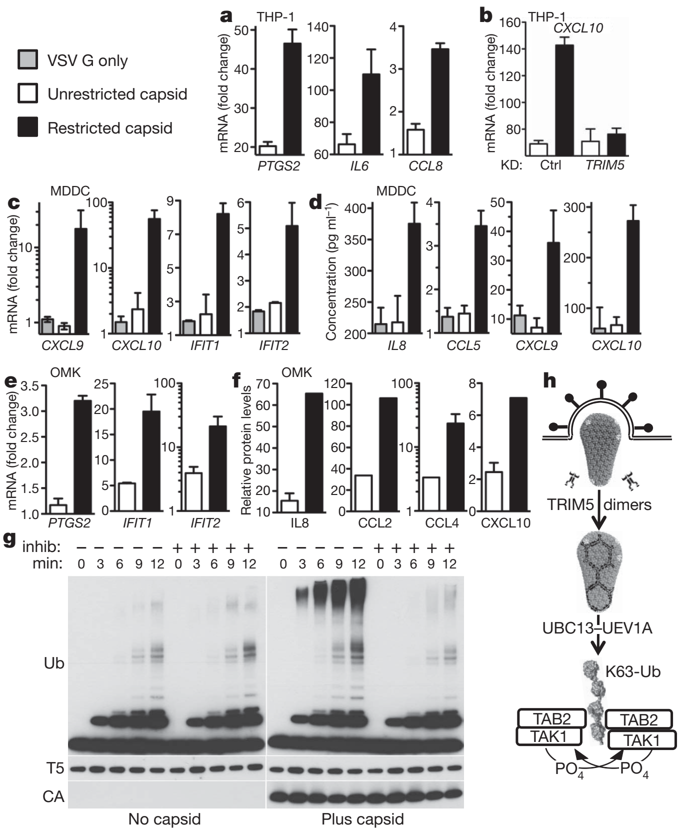


Figure 4 | Retrovirus capsid sensing by TRIM5. THP-1 cells (**a** and **b**), MDDCs (**c** and **d**), or owl monkey kidney cells (OMK; **e** and **f**), were challenged with matched pairs of VSV G-pseudotyped particles bearing retrovirion capsids that are restricted by the TRIM5 orthologue endogenous to that cell type (black bars), or unrestricted (white bars), or VSV G-derived particles that are devoid of capsid (grey bars). Restricted capsids were from N-tropic MLV (**a**–**d**) or HIV-1 (**e** and **f**). Unrestricted capsids were B-tropic MLV (**a** and **b**), N/B-tropic MLV (**c** and **d**) or SIV_{MAC239} (**e** and **f**). Particles bore viral genomes in **a**, **b**, **e** and **f**, but not in **c** and **d**. mRNA was harvested for qRT–PCR (**a**–**c** and **e**) and reported as fold change versus media control. Protein in the supernatant was quantified (**d** and **f**). Bars show means ± s.d. ($n = 3$), and are representative of at least three independent experiments. **g**, Immunoblots with the indicated antibodies of products from *in vitro* time-course with ATP, Ub, UBE1 (E1), UBC13–UEV1A (E2) and purified owl monkey TRIM5Cyp, with or without assembled HIV-1 capsid-A14C/E45C, and with or without competitive inhibitor MeIle4CsA. (**h**) Schematic showing entry of an HIV-1 virion core[27] (courtesy of Pornillos and Yeager) into the target cell cytoplasm where it induces dimeric TRIM5 to form a hexameric lattice[25] with increased E3 Ub ligase activity. With UBC13–UEV1A, TRIM5 synthesizes free K63 Ub chains that are recognized by TAB2, which multimerizes and activates the TAK1 kinase complex.

secretion of IL8, CCL2, CCL4 and CXCL10 proteins, was higher after challenge with the restricted virus (Fig. 4e, f).

TRIM5 senses retrovirus capsids in the target cell cytoplasm (Fig. 4a–f) and activates MAPK- and NF-κB-dependent transcription via the synthesis of TAK1-activating, K63-linked Ub chains (Fig. 3g–j). If these observations were linked functionally, interaction with capsid would be expected to stimulate the synthesis of Ub chains by TRIM5Cyp. Soluble, recombinant HIV-1 capsid or capsid hexamers generated by the oxidation of recombinant capsid bearing strategically-placed cysteine substitutions (A14C/E45C/W184A/M185A)[23] had no effect on the synthesis of K63-linked Ub chains (data not shown). Current models of the HIV-1 capsid lattice are based on cylinders generated under high salt with either capsid or capsid-nucleocapsid fusion protein[17,24]; both preparations were generated but the high salt necessary to maintain capsid cylinders blocked E3 Ub ligase activity. Capsid cylinders were then assembled with A14C/E45C-substituted

capsid protein in 1 M NaCl and the cysteines were oxidized. These oxidized cylinders were stable in the absence of salt (Supplementary Fig. 12a) and greatly stimulated the production of K63-linked Ub chains by TRIM5Cyp (Fig. 4g). No Ub-linked products were detected with anti-capsid (p24) or anti-TRIM5 antibodies, indicating that the reaction products were unattached Ub chains (Fig. 4g).

Finally, two factors that disrupt the HIV-1 capsid-TRIM5Cyp interaction and block restriction activity—a non-immunosuppressive cyclosporine analogue[1] or the TRIM5Cyp(H436Q) mutant protein[13]—each eliminated the enhancement of E3 Ub ligase activity by the A14C/E45C capsid cylinders, without effect on the baseline activity in the absence of capsid (Fig. 4g and Supplementary Figs 7d and 12b–d).

The experiments presented here demonstrate that TRIM5 is a multifunctional component of the innate immune system. In addition to functioning as a retroviral capsid-specific restriction factor, TRIM5 synthesizes K63 Ub chains that activate TAK1 and inflammatory transcription, most probably via multimerization of the TAK1-associated Ub-binding protein TAB2 (ref. 11; Fig. 4h). This activity was greatly increased by the hexameric capsid lattice, a molecular signature of HIV-1 and other retroviruses. TRIM5, then, satisfies criteria for a bona fide PRR[10]. Interestingly, TRIM5 spontaneously forms an hexagonal lattice that is complementary to the capsid lattice[25], but the efficiency of TRIM5 lattice formation is greatly stimulated by the capsid hexameric lattice[25] (Fig. 4h). Little is known about how the innate immune system detects retroviruses[26] and the discovery that TRIM5 acts as a PRR is an important step towards filling this critical gap. The cellular factors required for TRIM5 E3 activity and inflammatory gene induction, UBC13, UEV1A and TAK1, also promoted capsid-specific restriction activity, indicating that the multiple functions of TRIM5 are mechanistically linked. Identification of relevant TAK1 substrates will inform future attempts to pinpoint the mechanism of restriction.

METHODS SUMMARY

Plasmids, cells and viruses. These methods were described previously[1,13,19] or are detailed in the Supplementary Information.

Recombinant protein. Production of full-length, soluble, TRIM5Cyp is described in the supplement. CA A14C E45C was produced and assembled into tubes as described[23,24].

Microarray. Illumina HumanHT-12 V3.0 expression bead chips were probed with RNA from *TRIM5* knockdown THP-1 cells. Data set and methods are available at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession number GSE25041.

Received 31 July 2010; accepted 3 March 2011.

1. Sayah, D. M., Sokolskaja, E., Berthoux, L. & Luban, J. Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* **430,** 569–573 (2004).
2. Stremlau, M. *et al.* The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature* **427,** 848–853 (2004).
3. Sebastian, S. & Luban, J. TRIM5α selectively binds a restriction-sensitive retroviral capsid. *Retrovirology* **2,** 40 (2005).
4. Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5α restriction factor. *Proc. Natl Acad. Sci. USA* **103,** 5514–5519 (2006).
5. Berthoux, L. *et al.* As_2O_3 enhances retroviral reverse transcription and counteracts Ref1 antiviral activity. *J. Virol.* **77,** 3167–3180 (2003).
6. Shi, M. *et al.* TRIM30α negatively regulates TLR-mediated NF-κB activation by targeting TAB2 and TAB3 for degradation. *Nature Immunol.* **9,** 369–377 (2008).
7. Tareen, S. U. & Emerman, M. Human Trim5α has additional activities that are uncoupled from retroviral capsid recognition. *Virology* **409,** 113–120 (2011).
8. Luban, J. & Cyclophilin, A. TRIM5, and resistance to human immunodeficiency virus type 1 infection. *J. Virol.* **81,** 1054–1061 (2007).
9. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-β enhanceosome. *Cell* **129,** 1111–1123 (2007).
10. Ishii, K. J., Koyama, S., Nakagawa, A., Coban, C. & Akira, S. Host innate immune receptors and beyond: making sense of microbial infections. *Cell Host Microbe* **3,** 352–363 (2008).
11. Xia, Z. P. *et al.* Direct activation of protein kinases by unanchored polyubiquitin chains. *Nature* **461,** 114–119 (2009).
12. Kornbluth, R. S., Oh, P. S., Munis, J. R., Cleveland, P. H. & Richman, D. D. Interferons and bacterial lipopolysaccharide protect macrophages from productive infection by human immunodeficiency virus *in vitro*. *J. Exp. Med.* **169,** 1137–1151 (1989).
13. Neagu, M. R. *et al.* Potent inhibition of HIV-1 by TRIM5-cyclophilin fusion proteins engineered from human components. *J. Clin. Invest.* **119,** 3035–3047 (2009).
14. Roe, T., Reynolds, T. C., Yu, G. & Brown, P. O. Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* **12,** 2099–2108 (1993).
15. Berthoux, L., Sebastian, S., Sokolskaja, E. & Luban, J. Cyclophilin A is required for TRIM5α-mediated resistance to HIV-1 in Old World monkey cells. *Proc. Natl Acad. Sci. USA* **102,** 14849–14853 (2005).
16. Zeng, W. *et al.* Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* **141,** 315–330 (2010).
17. Langelier, C. R. *et al.* Biochemical characterization of a recombinant TRIM5α protein that restricts human immunodeficiency virus type 1 replication. *J. Virol.* **82,** 11682–11694 (2008).
18. Yin, Q., Lamothe, B., Darnay, B. G. & Wu, H. Structural basis for the lack of E2 interaction in the RING domain of TRAF2. *Biochemistry* **48,** 10558–10567 (2009).
19. Sokolskaja, E., Berthoux, L. & Luban, J. Cyclophilin A and TRIM5α independently regulate human immunodeficiency virus type 1 infectivity in human cells. *J. Virol.* **80,** 2855–2862 (2006).
20. Yan, N., Regalado-Magdos, A. D., Stiggelbout, B., Lee-Kirsch, M. A. & Lieberman, J. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nature Immunol.* **11,** 1005–1013 (2010).
21. Perron, M. J. *et al.* TRIM5α mediates the postentry block to N-tropic murine leukemia viruses in human cells. *Proc. Natl Acad. Sci. USA* **101,** 11827–11832 (2004).
22. Ulm, J. W., Perron, M., Sodroski, J. & Mulligan, R. C. Complex determinants within the Moloney murine leukemia virus capsid modulate susceptibility of the virus to Fv1 and Ref1-mediated restriction. *Virology* **363,** 245–255 (2007).
23. Pornillos, O. *et al.* X-ray structures of the hexameric building block of the HIV capsid. *Cell* **137,** 1282–1292 (2009).
24. Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T. & Sundquist, W. I. Assembly and analysis of conical models for the HIV-1 core. *Science* **283,** 80–83 (1999).
25. Ganser-Pornillos, B. K. *et al.* Hexagonal assembly of a restricting TRIM5α protein. *Proc. Natl Acad. Sci. USA* **108,** 534–539 (2011).
26. Medzhitov, R. & Littman, D. HIV immunology needs a new direction. *Nature* **455,** 591 (2008).
27. Pornillos, O., Ganser-Pornillos, B. K. & Yeager, M. Atomic-level modelling of the HIV capsid. *Nature* **469,** 424–427 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Baltimore, M. J. Birrer, J. Brojatsch, A. Cimarelli, A. Delaco, S. Elledge, M. Emerman, W. Ferlin, D. Garcin, S. Ghosh, O. Haller, T. Hatziioannou, J. Hiscott, A. Iwasaki, D. Kolakofsky, M. Kosco-Vilbois, H. Malik, R. Medzhitov, M. R. Neagu, G. Napolitani, P. Palese, D. Pinschewer, O. Pornillos, L. Roux, O. Schwartz, M. Strubin, V. Studer, W. Sundquist, G. Towers, D. Trono, J. Tschopp, M. Yeager, M. Zufferey, and the Functional Genomics Center (Zürich), for ideas, technical assistance, and reagents. This work was supported by NIH grant RO1AI59159 to J.L., NIH grant R21AI087467 to W.M., Swiss National Science Foundation grant 3100A0-128655 to J.L. and 3100A0-122342 to M.G. and UZH Forschungskredit 54041402 to S.Z.

Author Contributions T.P., S.H., J.G., C.R., C.S., M.P., W.M., M.G.G. and J.L. designed the experiments; T.P., S.H., D.M., S.Z., J.G., J.La., C.R., F.A.S., M.P., A.B., P.D.U. and L.C. performed the experiments. All authors contributed to the assembly and writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.L. (jeremy.luban@unige.ch).

doi:10.1038/nature09860

A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK

Damian F. Brennan[1]*, Arvin C. Dar[2]*, Nicholas T. Hertz[2], William C. H. Chao[1], Alma L. Burlingame[3], Kevan M. Shokat[2] & David Barford[1]

In metazoans, the Ras–Raf–MEK (mitogen-activated protein-kinase kinase)–ERK (extracellular signal-regulated kinase) signalling pathway relays extracellular stimuli to elicit changes in cellular function and gene expression. Aberrant activation of this pathway through oncogenic mutations is responsible for a large proportion of human cancer. Kinase suppressor of Ras (KSR)[1–3] functions as an essential scaffolding protein to coordinate the assembly of Raf–MEK–ERK complexes[4,5]. Here we integrate structural and biochemical studies to understand how KSR promotes stimulatory Raf phosphorylation of MEK (refs 6, 7). We show, from the crystal structure of the kinase domain of human KSR2 (KSR2(KD)) in complex with rabbit MEK1, that interactions between KSR2(KD) and MEK1 are mediated by their respective activation segments and C-lobe αG helices. Analogous to BRAF (refs 8, 9), KSR2 self-associates through a side-to-side interface involving Arg 718, a residue identified in a genetic screen as a suppressor of Ras signalling[1–3]. ATP is bound to the KSR2(KD) catalytic site, and we demonstrate KSR2 kinase activity towards MEK1 by *in vitro* assays and chemical genetics. In the KSR2(KD)–MEK1 complex, the activation segments of both kinases are mutually constrained, and KSR2 adopts an inactive conformation. BRAF allosterically stimulates the kinase activity of KSR2, which is dependent on formation of a side-to-side KSR2–BRAF heterodimer. Furthermore, KSR2–BRAF heterodimerization results in an increase of BRAF-induced MEK phosphorylation via the KSR2-mediated relay of a signal from BRAF to release the activation segment of MEK for phosphorylation. We propose that KSR interacts with a regulatory Raf molecule in *cis* to induce a conformational switch of MEK, facilitating MEK's phosphorylation by a separate catalytic Raf molecule in *trans*.

To understand how KSR regulates Raf-dependent MEK activation[6,7], we determined the crystal structure of the KSR2(KD)–MEK1 complex (Fig. 1 and Supplementary Table 1). KSR2(KD)–MEK1 dimers assemble into a tetramer through a KSR2(KD) homodimer interface centred on Arg 718 (Supplementary Fig. 1a). In solution, KSR2(KD)–MEK1 tetramers and dimers exist in dynamic exchange, indicating a relatively weak KSR2(KD) homodimer interaction (Supplementary Fig. 1b). KSR2(KD) and MEK1 molecules interact with their catalytic sites facing each other through their activation segments and αG helices (Fig. 1 and Supplementary Fig. 2a), reminiscent of MEK[10] and Chk2 (ref. 11) homodimers (Supplementary Fig. 2b). MEK's activation segment, incorporating Raf phosphorylation sites Ser 218^{M} and Ser 222^{M} (where superscript M indicates MEK1), comprises a short α-helix connected to a segment of extended chain that forms an antiparallel β-sheet with the KSR2 activation segment (Fig. 1b). This mutually constrained interaction of the MEK1 and KSR2 activation segments is stabilized by contacts between non-polar residues. The structural similarity of MEK1 when in complex with KSR2 to isolated MEK1, including those bound to allosteric inhibitors[10,12,13] (Supplementary Fig. 3), indicates that MEK inhibitors engage a physiologically relevant conformation of the kinase.

MEK and KSR form constitutive complexes[14,15] stable to Raf phosphorylation[16,17], even though Ser 218^{M} and Ser 222^{M} phosphorylation would alter the conformation of the MEK1 activation segment. The integrity of the complex is probably conferred by the more extensive interface created by engagement of their respective αG helices (Fig. 1a, c). This is consistent with studies showing that mutation of MEK within a conserved hydrophobic motif (Met 308^{M} to Ile 310^{M}), contiguous with the αG helix, disrupts MEK–KSR1 interactions[17]. In the

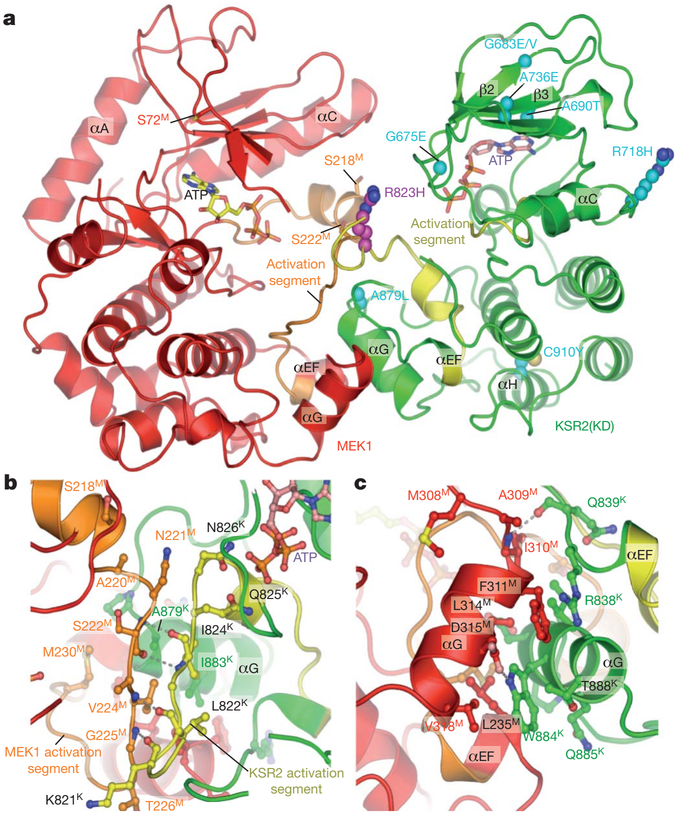


Figure 1 | Structure of the KSR2(KD)–MEK1 heterodimer. **a**, Overall view of the complex showing the face-to-face configuration of KSR2 and MEK1. KSR loss-of-function mutations[1–3] are shown with C atoms in cyan. A lung adenocarcinoma-associated mutation of KSR2 (R823H)[28] is located on the KSR2 activation segment interacting with the MEK1 activation segment. Sites of MEK1 phosphorylation (Ser 72, 218 and 222) are indicated. **b**, Details of the β-sheet formed between the activation segments of MEK1 and KSR2. **c**, Details of αG helix contacts.

[1]Section of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, UK. [2]Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, California 94107, USA. [3]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94107, USA.
*These authors contributed equally to this work.

complex, Ala 309$^{\mathrm{M}}$ and Ile 310$^{\mathrm{M}}$ contact KSR2 directly, whereas Met 308$^{\mathrm{M}}$ helps define the conformation of the αG helix (Fig. 1c). Furthermore, the KSR2(KD)–MEK1 structure rationalizes KSR genetic screens[2,16,18]. Tyr substitution of Cys 910$^{\mathrm{K}}$ (where superscript K indicates KSR2), a buried residue of the C-lobe αH helix, would destabilize the C-lobe, altering the αG helix conformation (Fig. 1a and Supplementary Fig. 4). In addition, replacement of Ala 879$^{\mathrm{K}}$ within the KSR2 αG helix with a bulky Leu residue in a KSR loss-of-function mutant[1] is not readily accommodated at the KSR–MEK interface (Fig. 1a, b), consistent with our finding that a KSR2(KD) A879L mutant destabilizes MEK1–KSR2(KD) interactions (Supplementary Fig. 5).

The KSR2(KD)–MEK1 crystal structure indicated that KSR has the potential for catalytic activity[19] possibly necessary for KSR function[6,20]. Electron density maps reveal well-defined density for ATP-Mg^{2+} at the KSR2(KD) catalytic site (Fig. 2a). Although Asp 803$^{\mathrm{K}}$ of the DFG motif is shifted slightly out of position, the metal-coordinating Asn 791$^{\mathrm{K}}$ and the general base Asp 786$^{\mathrm{K}}$ of the catalytic loop adopt conformations typical of conventional protein kinases (Fig. 2b). However, owing to the inactive position of its αC helix, reminiscent of inactive c-Src(KD) (refs 9, 21), KSR2(KD) adopts an inactive conformation (Supplementary Fig. 6). Conservation of ATP-binding-site residues, and the mapping of KSR loss-of-function mutants to the ATP pocket[2,3], indicate that ATP binding is necessary for KSR function (Fig. 1a and Supplementary Figs 4 and 7). A Thr substitution for Ala 690$^{\mathrm{K}}$ of β3, for example, would impede ATP binding, whereas a Glu substitution of Gly 675$^{\mathrm{K}}$ within the Gly-rich loop would disrupt the phosphate-binding site. Whereas *Drosophila* and *Caenorhabditis elegans* KSR share the same catalytic residues as conventional protein kinases, substitution of Arg for the conserved ATP-coordinating Lys residue in mammalian KSRs, the lack of reproducible kinase catalytic activity, and the capacity of kinase-impaired mutants of KSR to mediate MEK phosphorylation and MAP kinase signalling[16,22], have implicated KSR as a pseudokinase. Studies of pseudokinases such as HER3 indicate that functions other than catalysis are critical for their role in signal transduction cascades, and that their extremely low kinase activity may

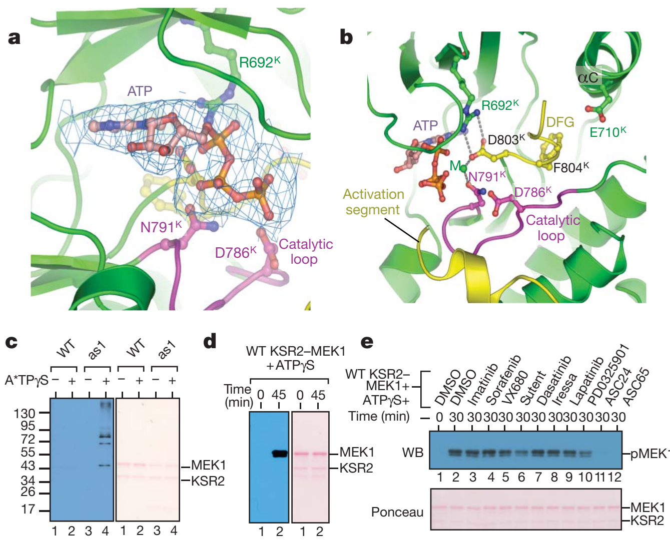


Figure 2 | KSR2 is a protein kinase. **a**, $2|F_o| - |F_c|$ electron density map for ATP bound to KSR2. **b**, Details of the catalytic site of KSR2. M, metal ion. **c**, Wild-type (WT) and T739G KSR (as1) KSR2–MEK1 complexes assayed in the presence of A*TPγS. Molecular mass markers in kDa are indicated along the left side of **c**. Phosphorylated proteins detected by western blot (WB) analysis (left) and total protein stained using ponceau red (right) are shown. **d**, Phosphorylation of MEK1 within wild-type KSR2–MEK1 combined with ATPγS. Phospho- (left) and total (right) proteins detected as in **c**. **e**, Enzymatic assay as in **d** in the presence of 10 μM of the listed inhibitors. Phospho-MEK1 is shown.

be sufficient to catalyse physiologically relevant transphosphorylation within a protein complex[23,24].

To examine putative KSR2 catalytic function we conducted *in vitro* kinase assays using an analogue-specific (as1) mutant of KSR2 (T739$^{\mathrm{K}}$G) incubated with the modified ATP analogue N6-phenethyl ATPγS (A*TPγS). Resultant phosphorylation products include MEK1, specifically within the as1-KSR2(KD)–MEK1 complex (Fig. 2c, lane 4). MEK1 within the wild-type KSR2(KD)–MEK1 complex incubated with unmodified ATPγS was also reproducibly phosphorylated (Fig. 2d), an activity attributed to KSR2 because purified MEK1 does not autophosphorylate (Supplementary Fig. 8a).

To exclude further the possibility of MEK1 phosphorylation resulting from MEK1 autophosphorylation or by a contaminating kinase, we examined the influence of inhibitors in the KSR2(KD)–MEK1 assay (Fig. 2e and Supplementary Table 2). None of the compounds tested blocked 100% of the MEK1 phosphorylation reaction, including sorafenib, a Raf inhibitor, and PD0325901, a MEK inhibitor. Although the MEK1 inhibitors Sutent and PD0325901 blocked 50% of the reaction, other MEK1 inhibitors (VX680 and dasatinib) showed no inhibition. It is possible that Sutent and PD0325901 are binding to MEK1 and influencing its phosphorylation, an established feature of PD0325901 (ref. 25). The pan-protein kinase inhibitors ASC24 and ASC65 (ref. 26) block 100% of the activity in the KSR2(KD)–MEK1 kinase complex, indicating that KSR2 is a druggable target (Fig. 2e).

We next sought to identify the site(s) of phosphorylation in MEK1 catalysed by KSR2. We did not detect reactivity from activation segment phospho-specific antibodies (Ser 222$^{\mathrm{M}}$ and Ser 218$^{\mathrm{M}}$/Ser 222$^{\mathrm{M}}$) or for other MEK1 sites including Thr 286$^{\mathrm{M}}$ or Ser 298$^{\mathrm{M}}$ (data not shown). Using mass spectrometry, we identified Ser 24$^{\mathrm{M}}$ and Ser 72$^{\mathrm{M}}$ as KSR2-dependent phosphorylation sites in MEK1 not previously assigned to any known kinase (Supplementary Fig. 9a, b). KSR2 phosphorylates MEK1 with low discrimination because replacing Ser 24$^{\mathrm{M}}$ and Ser 72$^{\mathrm{M}}$ with alanines only reduced overall MEK1 phosphorylation by ~10% (Supplementary Fig. 8b). Additional KSR2-dependent sites were identified as Ser 18$^{\mathrm{M}}$ and Thr 23$^{\mathrm{M}}$ (Supplementary Fig. 9c). The stoichiometry of KSR2-catalysed MEK1 phosphorylation is 1.25% (data not shown), indicating a very inefficient reaction.

KSR assembles MEK–KSR–Raf ternary complexes responsible for promoting Raf phosphorylation of MEK[8,17,27] in which Raf interacts with the CA1 domain and CA5-kinase domain of KSR (refs 17, 27). The KSR2(KD) homodimer interface has a striking structural resemblance to the homodimer interface of BRAF(KD) (ref. 9), termed the side-to-side dimer interface and proposed to mediate Raf(KD)–KSR(KD) heterodimer interactions[8] (Fig. 3a, b). Both interfaces are centred on a conserved N-lobe Arg residue (Arg 718$^{\mathrm{K}}$ of KSR2 and Arg 509 of BRAF). Arg 718$^{\mathrm{K}}$ is functionally significant because its replacement with His, a loss-of-function mutation[2] (Fig. 1a), abolishes the capacity of Raf to activate MEK[8,22]. Notably, His substitution of Arg, shown to dissociate BRAF(KD) (ref. 8) and KSR2(KD) homodimers (Supplementary Fig. 10), would not be compatible with a Raf(KD)–KSR(KD) heterodimer. KSR2(KD) and BRAF(KD) homodimers adopt different quaternary modes (termed I and A, respectively), and only a subset of intersubunit contacts are common to KSR2(KD) and BRAF(KD) side-to-side dimer interfaces (Fig. 3 and Supplementary Figs 7 and 11). Raf phosphorylation of MEK is promoted by enforced dimerization of Raf and KSR kinase domains through the side-to-side dimer interface[8]. Conservation in KSR of BRAF side-to-side contact residues indicates that KSR and BRAF form heterodimers by means of the BRAF quaternary A mode.

The quaternary structure of the KSR2(KD)–BRAF(KD) heterodimer determines the tertiary structure of the KSR2 αC helix. An inactive conformation of the αC helix is incompatible with the BRAF quaternary A mode. Thus, formation of a KSR2–BRAF heterodimer (A mode) would be accompanied by a shift of the KSR2 αC helix to an active conformation (Supplementary Fig. 12), rationalizing the conservation of Phe 707$^{\mathrm{K}}$ and Glu 710$^{\mathrm{K}}$ of αC (Supplementary Figs 7 and 13). A

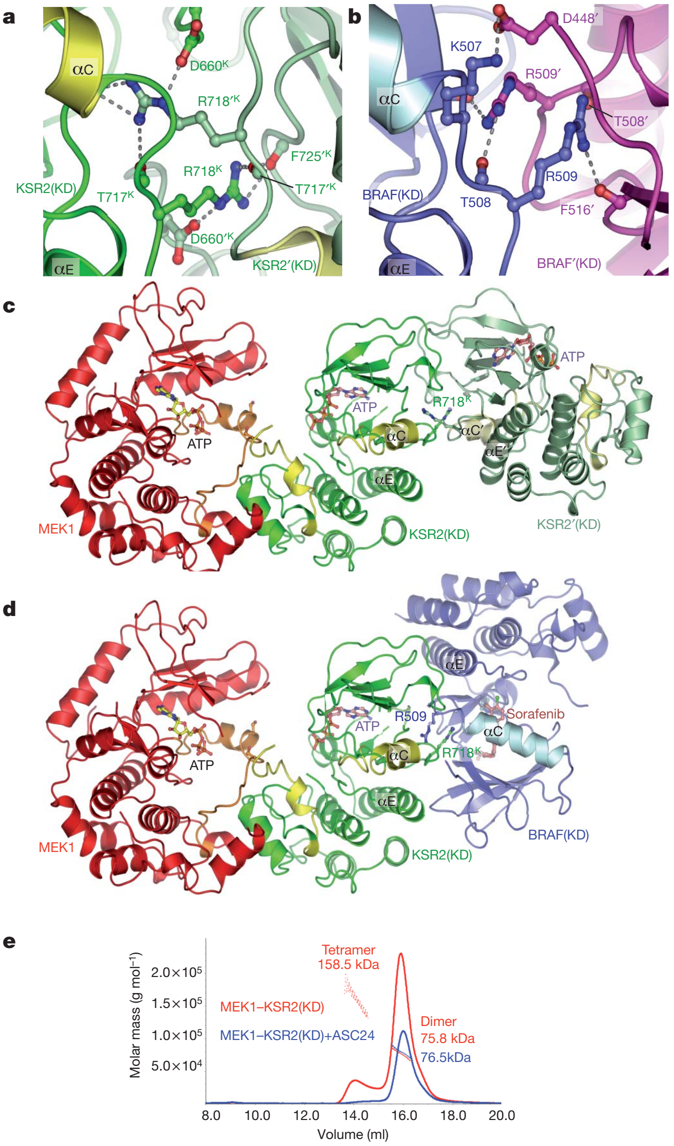


Figure 3 | KSR2 and BRAF homodimerize through a conserved side-to-side interface, but generate different quaternary structures. a, Details of the KSR2–KSR2 interface centred on Arg 718. Prime (′) denotes residues from the opposite subunit (quaternary mode I). **b**, Details of the BRAF–BRAF interface centred on Arg 509 (quaternary mode A) (Protein Data Bank 1uwh)[9]. **c**, MEK1–KSR2–KSR2 ternary complex. View as in Fig. 1a. **d**, Predicted MEK1–KSR2–BRAF ternary complex (quaternary mode A). **e**, Multiple angle laser light scattering experiment showing that ASC24 dissociates the KSR2(KD)–MEK1 heterotetramer into a heterodimer. The lower peak height in the presence of ASC24 is due to ASC24-induced protein precipitation.

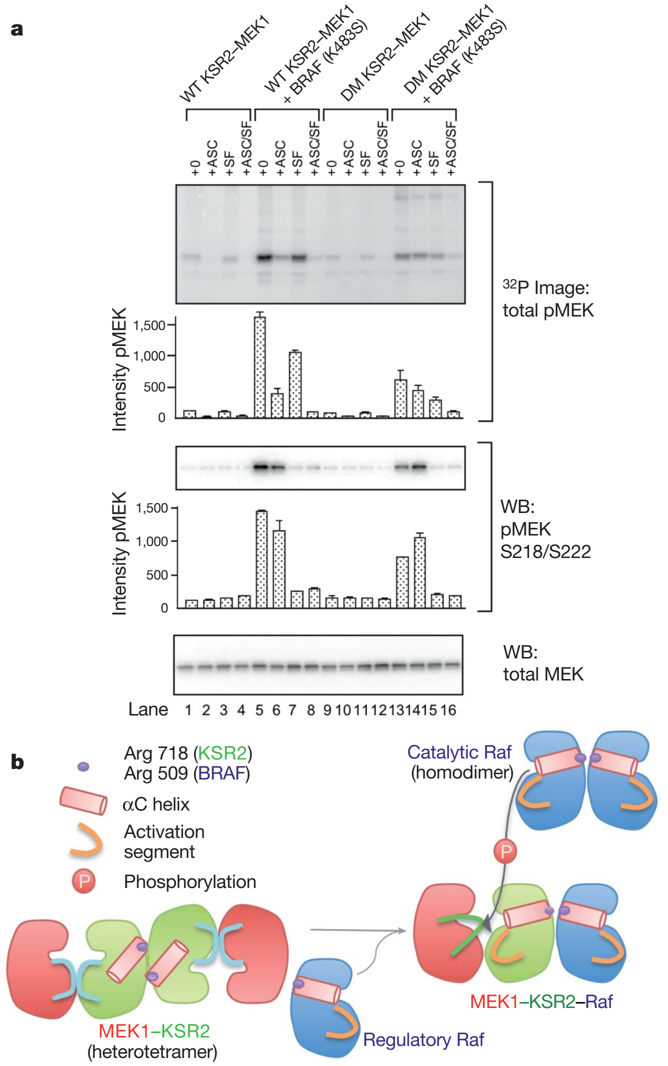


Figure 4 | Stimulation of MEK phosphorylation by an allosteric transition of KSR mediated by BRAF binding to the Arg 718 side-to-side interface. a, Biochemical reconstitution of KSR–Raf dimerization assessed based on MEK activation. Total phospho-MEK (top), S218/S222 site-specific phospho-MEK (middle), and total MEK (bottom; loading control) are shown. Bands were quantified and phospho-MEK levels were normalized relative to lane 1 in the top and middle panels, respectively. Error bars represent the mean ± s.d. ($n = 2$). **b**, Schematic of Raf-mediated activation of Raf phosphorylation of MEK. The side-to-side KSR2–Raf heterodimer (mode A) and Raf–Raf homodimer (mode A) are centred on Arg 718 (KSR2) and Arg 509 (BRAF). In the KSR2–MEK1 heterotetramer (left), the inaccessible activation segment of MEK1 is released through the interaction of KSR2 with a regulatory Raf molecule, induced by a conformational change of the αC helix, allowing catalytic Raf to phosphorylate MEK (right). The regulatory Raf molecule could be a catalytically impaired oncogenic mutant BRAF, whereas the catalytic Raf could also be Raf of a MEK1–KSR2–Raf ternary complex. Regulatory and catalytic Raf could be different isoforms (B-RAF or C-RAF). A Raf molecule bound to KSR through a side-to-side heterodimer interface is sterically unable to phosphorylate MEK in *cis*, an activity that therefore must be performed by a different Raf molecule.

BRAF-induced conformational shift of the KSR2 αC helix into the active position would probably enhance KSR2 kinase activity and influence its interaction with the activation segment of MEK, thereby affecting Raf's capacity to phosphorylate MEK. To test whether KSR2–BRAF heterodimerization stimulates KSR2 and BRAF-dependent MEK phosphorylation, we performed *in vitro* kinase assays. To overcome the weak and transient KSR2–BRAF interactions we immobilized His_6-tagged MEK1, KSR2 and BRAF onto Co^{2+} resin. As shown in Fig. 4a, the addition of kinase-impaired BRAF(K483S) to KSR2(KD)–MEK1 increases total MEK1 phosphorylation 15-fold (compare lanes 1 and 5). Interestingly, most of the increased activity comes from accelerated KSR2 catalytic function (as opposed to activity from additional BRAF). Approximately 70% of total MEK1 phosphorylation and less than 10% of S218/S222-specific activity was blocked by the KSR2 inhibitor ASC24 (Fig. 4a, lanes 5 and 6). Furthermore, most of the stimulated phosphorylation occurred at non-BRAF sites, as sorafenib only blocked 30% of total MEK1 phosphorylation but over 90% of the S218/S222-specific signal (Fig. 4a, lanes 5 and 7). Thus, KSR2 is the major kinase responsible for stimulated MEK1 phosphorylation in lane 5, and this occurs through a BRAF(K483S)-induced increase in KSR2 catalytic activity. That the R718H dimer interface mutation in KSR2 (KSR2(DM)–MEK) significantly reduced BRAF-induced stimulation indicates that KSR2 modulation by BRAF(K483S) depends primarily on the Arg 718-mediated

side-to-side interface of KSR2–BRAF heterodimers (Fig. 4a, compare lanes 5 and 13). In the absence of BRAF, the dimer mutation in KSR2 is functionally silent (Fig. 4a, compare lanes 1–4 and 9–12), indicating that wild-type and dimer mutant KSR2 are indistinguishable in a basal state.

Interestingly, BRAF(K483S) retained kinase activity (Supplementary Fig. 14); however, its intrinsic catalytic activity was not modulated through heterodimerization with KSR2. Indeed, with KSR2 inhibited by ASC24, MEK1 phosphorylation by BRAF(K483S) in wild-type and dimer mutant KSR2(KD)–MEK1 was approximately equal (Fig. 4a, lanes 6 and 14; middle panel). Thus, the twofold increase in MEK1 phosphorylation at the BRAF-specific sites, which were upregulated through KSR2–BRAF heterodimerization (lanes 5 and 13; middle panel), more likely resulted from a change in MEK1 structure coupled to allostery in KSR2 rather than an increase in intrinsic BRAF catalytic activity. KSR2(KD) heterodimerization with wild-type BRAF also stimulates BRAF-mediated MEK1 phosphorylation (Supplementary Fig. 15).

Control experiments excluded the possibility that either MEK1 autophosphorylation or contaminating kinases contribute to MEK1 phosphorylation (Supplementary Fig. 16). Thus, our data provide evidence for KSR2 as a protein kinase, stimulated by its interactions with BRAF. The physiological significance of KSR2 kinase activity awaits identification of its cellular substrate(s); however, its BRAF-triggered activation provides an *in vitro* barometer of a conformational change in KSR2. The kinase-independent function of KSR2 modulation of MEK1 is further borne out by the KSR2 inhibitor ASC24, which induces an unexpected increase in MEK1 phosphorylation by BRAF(K483S) (Fig. 4a, lanes 13 and 14; middle panel). Presumably, ASC24 induced a conformational change in KSR2 that partially mimics KSR Arg 718 side-to-side binding by BRAF, thus altering MEK1 and allowing for its increased phosphorylation, consistent with the capacity of ASC24 to disrupt the KSR2(KD) homodimer interface (Fig. 3e). Thus, a KSR inhibitor that binds to the KSR–MEK complex and antagonises the BRAF-induced allosteric switch could function analogously to the Ras suppressor: the KSR(R718H) mutation.

Here we reveal that KSR functions as an effector molecule, receiving a stimulatory signal from an activated regulatory Raf molecule and relaying this to modulate the accessibility of the MEK activation segment for phosphorylation by catalytic Raf (Fig. 4b). Our results highlight the complexity and adaptability of protein kinases, pseudokinases and scaffolding molecules to integrate signal transduction processes.

METHODS SUMMARY

Expression, purification and crystallization of KSR2–MEK1. *Homo sapiens* KSR2 kinase domain (KSR2(KD), residues 634–950) with an N-terminal His_6 tag was co-expressed with full-length MEK1 and the $p50^{Cdc37}$ Hsp90 co-chaperone in the baculovirus/Sf21 insect cell system. The KSR2(KD)–MEK1 complex was purified and crystallized as described in Methods. Crystal structure determination, mutagenesis, compound synthesis, enzyme assays, mass spectrometry and other procedures were performed as described in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 June 2010; accepted 20 January 2011.
Published online 27 March 2011.

1. Kornfeld, K., Hom, D. B. & Horvitz, H. R. The *ksr-1* gene encodes a novel protein kinase involved in Ras-mediated signaling in *C. elegans*. *Cell* **83,** 903–913 (1995).
2. Sundaram, M. & Han, M. The *C. elegans ksr-1* gene encodes a novel Raf-related kinase involved in Ras-mediated signal transduction. *Cell* **83,** 889–901 (1995).
3. Therrien, M. *et al.* KSR, a novel protein kinase required for RAS signal transduction. *Cell* **83,** 879–888 (1995).
4. Clapéron, A. & Therrien, M. KSR and CNK: two scaffolds regulating RAS-mediated RAF activation. *Oncogene* **26,** 3143–3158 (2007).
5. Kolch, W. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. *Nature Rev. Mol. Cell Biol.* **6,** 827–837 (2005).
6. Therrien, M., Michaud, N. R., Rubin, G. M. & Morrison, D. K. KSR modulates signal propagation within the MAPK cascade. *Genes Dev.* **10,** 2684–2695 (1996).
7. Michaud, N. R. *et al.* KSR stimulates Raf-1 activity in a kinase-independent manner. *Proc. Natl Acad. Sci. USA* **94,** 12792–12796 (1997).
8. Rajakulendran, T., Sahmi, M., Lefrancois, M., Sicheri, F. & Therrien, M. A dimerization-dependent mechanism drives RAF catalytic activation. *Nature* **461,** 542–545 (2009).
9. Wan, P. T. *et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116,** 855–867 (2004).
10. Ohren, J. F. *et al.* Structures of human MAP kinase kinase 1 (MEK1) and MEK2 describe novel noncompetitive kinase inhibition. *Nature Struct. Mol. Biol.* **11,** 1192–1197 (2004).
11. Cai, Z., Chehab, N. H. & Pavletich, N. P. Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. *Mol. Cell* **35,** 818–829 (2009).
12. Fischmann, T. O. *et al.* Crystal structures of MEK1 binary and ternary complexes with nucleotides and inhibitors. *Biochemistry* **48,** 2661–2674 (2009).
13. Iverson, C. *et al.* RDEA119/BAY 869766: a potent, selective, allosteric inhibitor of MEK1/2 for the treatment of cancer. *Cancer Res.* **69,** 6839–6847 (2009).
14. Denouel-Galy, A. *et al.* Murine Ksr interacts with MEK and inhibits Ras-induced transformation. *Curr. Biol.* **8,** 46–55 (1998).
15. Yu, W., Fantl, W. J., Harrowe, G. & Williams, L. T. Regulation of the MAP kinase pathway by mammalian Ksr through direct interaction with MEK and ERK. *Curr. Biol.* **8,** 56–64 (1998).
16. Stewart, S. *et al.* Kinase suppressor of Ras forms a multiprotein signaling complex and modulates MEK localization. *Mol. Cell. Biol.* **19,** 5523–5534 (1999).
17. McKay, M. M., Ritt, D. A. & Morrison, D. K. Signaling dynamics of the KSR1 scaffold complex. *Proc. Natl Acad. Sci. USA* **106,** 11022–11027 (2009).
18. Muller, J., Cacace, A. M., Lyons, W. E., McGill, C. B. & Morrison, D. K. Identification of B-KSR1, a novel brain-specific isoform of KSR1 that functions in neuronal signaling. *Mol. Cell. Biol.* **20,** 5529–5539 (2000).
19. Zhang, Y. *et al.* Kinase suppressor of Ras is ceramide-activated protein kinase. *Cell* **89,** 63–72 (1997).
20. Sugimoto, T., Stewart, S., Han, M. & Guan, K. L. The kinase suppressor of Ras (KSR) modulates growth factor and Ras signaling by uncoupling Elk-1 phosphorylation from MAP kinase activation. *EMBO J.* **17,** 1717–1727 (1998).
21. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385,** 595–602 (1997).
22. Douziech, M., Sahmi, M., Laberge, G. & Therrien, M. A. KSR/CNK complex mediated by HYP, a novel SAM domain-containing protein, regulates RAS-dependent RAF activation in *Drosophila*. *Genes Dev.* **20,** 807–819 (2006).
23. Jura, N., Shan, Y., Cao, X., Shaw, D. E. & Kuriyan, J. Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc. Natl Acad. Sci. USA* **106,** 21608–21613 (2009).
24. Shi, F., Telesco, S. E., Liu, Y., Radhakrishnan, R. & Lemmon, M. A. ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc. Natl Acad. Sci. USA* **107,** 7692–7697 (2010).
25. Alessi, D. R., Cuenda, A., Cohen, P., Dudley, D. T. & Saltiel, A. R. PD 098059 is a specific inhibitor of the activation of mitogen-activated protein kinase kinase *in vitro* and *in vivo*. *J. Biol. Chem.* **270,** 27489–27494 (1995).
26. Statsuk, A. V. *et al.* Tuning a three-component reaction for trapping kinase substrate complexes. *J. Am. Chem. Soc.* **130,** 17568–17574 (2008).
27. Roy, F., Laberge, G., Douziech, M., Ferland-McCollough, D. & Therrien, M. KSR is a scaffold required for activation of the ERK/MAPK module. *Genes Dev.* **16,** 427–438 (2002).
28. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446,** 153–158 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by a Cancer Research UK grant to D.B., ICR studentships to D.F.B. and W.C.H.C. and HHMI grant to K.M.S. We thank staff at the ESRF for help with data collection and K. Wood and V. Good for help with protein production and Z. Zhang for assistance with cloning. Mass spectrometry was made possible by NIH grants NCRR RR015804 and NCRR RR001614. The MEK1/$p50^{Cdc37}$ baculovirus was a gift from C. Vaughan. We thank Cell Signaling Technologies for help with phosphospecific MEK antibodies.

Author Contributions D.F.B. determined and analysed the MEK1–KSR2 structure; A.C.D. conducted biochemical analysis of Raf–KSR–MEK phosphorylation and inhibitor studies; N.T.H. carried out phosphoproteomics mass spectrometry studies; W.C.H.C. helped with protein production; A.L.B. analysed mass spectrometry data; K.M.S. designed and analysed experiments relating to the Raf–KSR–MEK phosphorylation and inhibitor studies; and D.B. designed experiments and analysed data.

Author Information Coordinates and structure factors have been deposited in the RCSB Protein Data Bank with accession numbers 2y4i and 2y4isf, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B. (david.barford@ic.ac.uk) or K.M.S. (shokat@cmp.ucsf.edu).

METHODS

Cloning, expression and purification of KSR2(KD)–MEK1. A variety of KSR isoforms were screened for expression and crystallization. *Homo sapiens* KSR2 kinase domain (KSR2(KD), residues 634–950) with an N-terminal His_6 tag was co-expressed with full-length rabbit MEK1 (with 3C Precission-cleavable His_6 tag) and the $p50^{Cdc37}$ Hsp90 co-chaperone in the baculovirus/Sf21 insect cell system (a gift from C. Vaughan). KSR2(KD)–MEK1 was purified using Co^{2+}-Talon resin (Clontech) and incubated with GST-tagged Rhinovirus 3C Precission and λ-phosphatase overnight. The protein was then applied to a RESOURCE S column (GE Healthcare) connected to a 5-ml GST column (GE Healthcare) to trap 3C protease. Free MEK1 was either present in the flow through or eluted early. The KSR2(KD)–MEK1 complex eluted as a later peak. Uncleaved protein and λ-phosphatase were removed using Talon resin and the KSR2(KD)–MEK1 complex was applied to a HiPrep 16/60 Superdex 200 gel filtration chromatography column. KSR2(KD)–MEK1 eluted as two peaks, which according to multiple angle light scattering (MALS) corresponded to a heterotetramer and a heterodimer. A smaller third peak is often seen that corresponded to MEK1. Sequential gel filtration has shown that the heterodimer and heterotetramer are dynamic species and so these peaks were pooled. Size exclusion chromatography was performed on mutant KSR2(KD)(R718H)–MEK1 complexes as for wild-type protein.

Crystallization of KSR2(KD)–MEK1. The complex was concentrated to 10 mg ml^{-1} and incubated with 5 mM Mg-ATP. Aggregation was cleared by filtration or by centrifugation for 20 min at 4 °C at 16,000*g*. Sitting-drop crystallization trials were set up using a Phoenix robot (Art Robbins Instruments) at 4 °C, 14 °C and 20 °C using a variety of commercial screens. Inclusion of 10 mM Pr acetate increased crystal size. Thousands of crystal growth conditions were trialled and in excess of 80 crystals were tested for diffraction. This led to a KSR2(KD)–MEK1 data set with Mg-ATP co-crystallization to 3.8 Å resolution from a crystal grown at 20 °C in 12% (w/v) PEG 3350, 0.2 M Na citrate, 10 mM Pr acetate, Bis-Tris propane pH 6.25. A data set to 3.45 Å was subsequently collected, which was also from a Mg-ATP co-crystallization in a similar condition with Mg acetate in place of Pr acetate. Crystallographic data were collected at beam line ID 14.4 ESRF, Grenoble and processed using MOSFLM and the CCP4 Program suite[29].

KSR2(KD)–MEK1 structure determination. BRAF shares ~30% sequence identity with the KSR kinase domain. A systematic approach to solving the structure by molecular replacement was undertaken. The CaspR automated molecular replacement server (http://www.igs.cnrs-mrs.fr/Caspr2/) was used[30] and is an automated molecular replacement pipeline that generates multiple sequence alignments using TCoffee[31], before homology model building with MODELLER[32], molecular replacement with AMoRe[29] and model refinement based on CNS[33]. Multiple CaspR runs were performed using all the potential space groups, with data scaled with a 5 Å high-resolution cutoff as well as a 3.8 Å resolution cutoff with both available structures of MEK (Protein Data Bank 1s9j and 2p55) as well as structures of BRAF (Protein Data Bank 1uwh and 3c4c) and with tyrosine kinases. This strategy led to a correct molecular replacement solution found using MEK and a MODELLER produced homology model of KSR based on the N-terminal lobe of BRAF and the C-terminal lobe of ephrin tyrosine kinase. Tyrosine kinases were suggested as search models despite having 15–20% sequence identity by GenTHREADER[34]. Molecular replacement was used to solve the 3.45 Å data set.

SeMet anomalous difference maps. To obtain independent phase information, we obtained anomalous diffraction data from SeMet incorporated KSR2(KD)–MEK1 crystals. SeMet-labelled KSR2(KD)–MEK1 was prepared in the insect cell system using a procedure modified from ref. 35. Peak data sets were collected for two crystals, although the best of these only diffracted to 7 Å. Anomalous difference maps were useful for confirming the location of Met residues and thus the chain trace of KSR2(KD).

Mutagenesis of BRAF(KD) and KSR2(KD)–MEK1. Point mutations of BRAF kinase-impaired (K483S), gate-keeper KSR2(KD)(T739G) and dimer mutant KSR2(KD)(R718H) were introduced using standard procedures. BRAF mutant and human $p50^{Cdc37}$ were cloned pairwise into a modified pFBDM vector (Z. Zhang). An N-terminally His_8-tagged human KSR2 kinase domain (634–950) was mutated to generate a dimerization-impaired mutant (R718H) in the modified pFBDM vector. In addition the following KSR2(KD) mutants were prepared: (1) dimer mutant KSR2(KD)(R718H); and (2) KSR2 kinase-impaired mutant (catalytic Asp) KSR2(KD)(D786A). Gene-containing pFBDM vectors were converted into bacmids for baculoviral productions. The kinase-dead MEK1 mutant was Lys 97 to Met, and a double phosphosite mutant Ser 24 and Ser 72 to Ala was prepared.

BRAF kinase domain purification. Viruses of BRAF mutants were used to infect Sf9 cells at a multiplicity of infection of 2. Cell pellets were harvested after 3 days of expression. BRAF was expressed and purified according to ref. 9 with minor modifications. Mutant proteins were purified with Talon metal ion affinity and ion-exchange chromatography.

MALS analysis of the KSR2(KD)–MEK1 complex. We used multiple angle laser light scattering at 658 nm using a Dawn Heleos light scattering instrument (Eldan) attached to a UV detector of purified KSR2(KD)–MEK1 complex that was run at room temperature on a 24 ml Superose 6 analytical gel filtration column (GE Healthcare) at a rate of 0.5 ml min^{-1} by HPLC (Varian).

MALS analysis of the KSR2(KD)–MEK1 + ASC24 complex. MALS gel filtration purified KSR2(KD)–MEK1 was concentrated to 2 mg ml^{-1} before multiple angle laser light scattering (MALS) experiments. ASC24 was incubated with KSR2(KD)–MEK1 at a final concentration of 4.5 mM at 4 °C, before the MALS experiment. Protein samples were run at 20 °C on a 24-ml capacity Superdex 200 analytical gel filtration column (GE Healthcare) at a rate of 0.5 ml min^{-1} by HPLC (Varian). MALS were performed on the eluate at 658 nm using a Dawn Heleos light scattering instrument (Eldan) attached to a UV detector.

***In vitro* kinase assays.** Wild-type and mutant forms of KSR2(KD)–MEK1 used for kinase assays were purified as described for the KSR2(KD)–MEK1 used for crystallization, except that the cleavage and gel filtration steps were omitted. Coomassie-blue-stained SDS–PAGE gels of purified proteins are shown in Supplementary Fig. 18.

Thiophosphorylation experiment. For the thiophosphorylation kinase assays, approximately 1 μM of purified KSR2–MEK1 complex (wild type or the as1 mutant) was pre-incubated for 10 min in reactions containing 10 μM of the indicated small molecule drugs, 10 mM Tris 7.8, 10 mM $MgCl_2$ and 2% DMSO. Reactions were initiated by adding ATPγS or N6-phenethyl ATPγS (A*TPγS) as indicated and allowed to proceed for 30 min, before being terminated with EDTA (0.025 mM final concentration). Thiophosphorylated reaction products were alkylated by adding 1.5 μl of a freshly prepared DMSO solution containing 50 mM para-nitrobenzyl mesylate (PBNM) to the mixtures. After incubation at room temperature for minimally 40 min, SDS-load dye was added, and samples were applied to a 4–20% Tris-HCl gel for SDS–PAGE separation. Proteins were transferred to nitrocellulose membrane, and then in sequence stained with ponceau dye, washed, blocked with 5% milk TBST, and then blotted using a phosphorothioate-specific monoclonal antibody (antibody 51-8; 1:10,000 in 5% milk)[36]. After removal of excess primary antibody and several washes, the membranes were incubated with HRP-conjugated anti-rabbit IgG secondary antibody (1:10,000 in 5% milk; Promega). Membranes were washed extensively, after which protein bands were visualized on film using enhanced chemiluminescence (Pierce).

Compound synthesis. Sorafenib, ASC24 and ASC65 were synthesized as previously described[26,37]. All other kinase inhibitors were obtained from Calbiochem or commercially available sources. Drugs were dissolved in DMSO before use.

Kinase assays on Co^{2+}-Talon beads. To overcome the weak and transient KSR2–BRAF interactions we immobilized His_6-tagged MEK1, KSR2 and BRAF onto Co^{2+}-Talon resin. We used this simplified reconstituted system to promote BRAF dimerization with purified KSR2–MEK1 complexes, and investigated the influence of these interactions on MEK1 phosphorylation. 1 μM of purified KSR2(KD)–MEK1 complex (wild type or mutants) was premixed with BRAF and then loaded onto Co^{2+}-Talon beads (Clontech). Beads were prepared from 50% slurry by washing in water. 2.0 μl of beads were used per reaction (25 μl total volume). For reactions containing free MEK1 (recovered), MEK1 isolated as a by-product from the KSR2(KD)–MEK1 complex was used. Free MEK1 (recovered) contains trace levels of KSR2 as detected by mass spectrometry. Kinases were diluted in buffer containing 10 mM Tris 7.8, 10 mM $MgCl_2$, 2% DMSO and 20 μM inhibitors as indicated. Reactions were initiated by the addition of 200 μM cold ATP supplemented with 5 μCi γ^{32}P-ATP and reactions were incubated at room temperature. After 90 min, reaction samples were mixed with SDS load buffer, heated, and electrophoresed on a 4–20% Tris-HCl SDS gradient gel. Gels were fixed, dried and exposed to a Typhoon phospho-imager to obtain the total MEK1 phosphorylation signal or transferred to nitrocellulose membranes and blotted with pMEK1/2 S217/S221 (Cell Signaling; 1:10,000 in 5% BSA TBST) or total MEK1 (Cell Signaling; 1:10,000 in 5% BSA TBST). Radioactive gels were scanned on a Typhoon imager (GE Healthcare) and bands were quantified using ImageQuant (GE Healthcare). Chemiluminescent signal on western blots was captured on a high-resolution digital camera using the AlphaInnotech system for quantification of pMEK on S218/S222 as shown in Fig. 4a. Graphing and analysis was performed using the program Prism.

Phosphopeptide mapping by tandem mass spectrometry. To identify phosphorylation sites we either enriched for phosphopeptides (ATP-treated KSR2(KD)–MEK1), or ran the whole protein digest directly on the mass spectrometer (A*TPγS-treated KSR2(KD)–MEK1). Preparations of the indicated kinase reaction were digested as described previously to retain the phosphorylation or thiophosphorylation mark[38]. The digested peptides were then desalted by using a C_{18} OMIX 100 μl Zip Tip. The phosphopeptides were then enriched by IMAC enrichment on a TiO_2 column as done previously[39]. The thiophosphorylated samples were not enriched. The peptides were run on a nano-LC system before

analysis on both a QSTAR Elie and a LTQ Orbitrap. For both, peptides were separated via a 60 min LC run (0–32% acetonitrile 0.1% formic acid). Separation was achieved by reversed-phase chromatography on a $75\,\mu m \times 15\,cm$ C18 column flowing at $350\,nl\,min^{-1}$ applied directly to a LTQ Orbitrap or a QSTAR Elite. For the LTQ Orbitrap, the three most intense ions above 10,000 counts were selected for subsequent fragmentation and MS analysis. Three CID spectra and three ETD spectra (200 = activation energy) were acquired. For the QSTAR Elite, the two most intense multiply charged peaks from each MS spectra were selected for fragmentation by CID, and MS analysis in the TOF mass analyser. For both, a dynamic exclusion window was applied that prevented the same peak from being selected for 1 min. Peak lists were generated by PAVA (in-house LTQ Orbitrap data) or Mascot (QSTAR Elite data) and analysed on Protein Prospector (LTQ Orbitrap: parent mass tolerance, 20 p.p.m., fragment mass tolerance, 0.6 Da, three missed cleavages, max four modifications, no constant modifications, and phospho serine/threonine variable modifications) (QSTAR Elite same as above except: parent mass tolerance, 200 p.p.m., fragment mass tolerance, 300 p.p.m., thiophospho serine/threonine variable modification).

29. Collaborative Computer Project 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
30. Claude, J. B., Suhre, K., Notredame, C., Claverie, J. M. & Abergel, C. CaspR: a web server for automated molecular replacement using homology modelling. *Nucleic Acids Res.* **32,** W606–W609 (2004).
31. Poirot, O., Suhre, K., Abergel, C., O'Toole, E. & Notredame, C. 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment, *Nucleic Acids Res.* **32,** W37–W40 (2004).
32. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815 (1993).
33. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
34. McGuffin, L. J. & Jones, D. T. Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* **19,** 874–881 (2003).
35. Cronin, C. N., Lim, K. B. & Rogers, J. Production of selenomethionyl-derivatized proteins in baculovirus-infected insect cells. *Protein Sci.* **16,** 2023–2029 (2007).
36. Allen, J. J. *et al.* A semisynthetic epitope for kinase substrates. *Nature Methods* **4,** 511–516 (2007).
37. Bankston, D. *et al.* A scaleable synthesis of BAY 43-9006: A potent raf kinase inhibitor for the treatment of cancer. *Org. Process Res. Dev.* **6,** 777–781 (2002).
38. Hertz, N. T. *et al.* Chemical genetic approach for kinase-substrate mapping by covalent capture of thiophosphopeptides and analysis by mass spectrometry. *Curr. Prot. Chem. Biol.* **2,** 15–36 (2010).
39. Trinidad, J. C. *et al.* Quantitative analysis of synaptic phosphorylation and protein expression. *Mol. Cell. Proteomics* **7,** 684–696 (2008).

LETTER

doi:10.1038/nature09872

Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development

Ikuhiro Okamoto[1]*, Catherine Patrat[1,2]*, Dominique Thépot[3], Nathalie Peynot[3], Patricia Fauque[4,5], Nathalie Daniel[3], Patricia Diabangouaya[1], Jean-Philippe Wolf[6], Jean-Paul Renard[3], Véronique Duranthon[3] & Edith Heard[1]

X-chromosome inactivation (XCI) in female mammals allows dosage compensation for X-linked gene products between the sexes[1]. The developmental regulation of this process has been extensively investigated in mice, where the X chromosome of paternal origin (Xp) is silenced during early embryogenesis owing to imprinted expression of the regulatory RNA, *Xist* (X-inactive specific transcript). Paternal XCI is reversed in the inner cell mass of the blastocyst and random XCI subsequently occurs in epiblast cells. Here we show that other eutherian mammals have very different strategies for initiating XCI. In rabbits and humans, the *Xist* homologue is not subject to imprinting and XCI begins later than in mice. Furthermore, *Xist* is upregulated on both X chromosomes in a high proportion of rabbit and human embryo cells, even in the inner cell mass. In rabbits, this triggers XCI on both X chromosomes in some cells. In humans, chromosome-wide XCI has not initiated even by the blastocyst stage, despite the upregulation of *XIST*. The choice of which X chromosome will finally become inactive thus occurs downstream of *Xist* upregulation in both rabbits and humans, unlike in mice. Our study demonstrates the remarkable diversity in XCI regulation and highlights differences between mammals in their requirement for dosage compensation during early embryogenesis.

Most eutherian mammals display random XCI in their soma[1], whereas marsupials show imprinted, paternal XCI[2]. In mice, however, XCI is also imprinted initially, with silencing of Xp from the 4–8-cell stage[3–6]. Xp remains inactive in the trophectoderm[7] but is reactivated in the inner cell mass[5,6] (ICM) with subsequent random XCI in epiblast cells. Initiation of XCI is controlled by the non-coding *Xist* transcript[8,9], which coats the chromosome in *cis* and triggers gene silencing. During murine pre-implantation development, *Xist* is maternally repressed and paternally expressed, consistent with its role in regulating imprinted XCI[3,4,6,9,10]. By the epiblast stage, the imprint is lost or ignored, Xp has been reactivated and *Xist* can be upregulated randomly, from either Xp or the maternally derived X chromosome, Xm. This second, random wave of XCI involves downregulation of pluripotency factors, activation of *Xist* by means of X-linked competence factors and precise, random monoallelic regulation of *Xist*, possibly through antisense transcription, homologous pairing and other events (for a review, see ref. 11). Mice are the only eutherian for which XCI has been analysed in detail during early embryogenesis. The fact that marsupials also show XCI and yet have no gene homologous to *Xist* (refs 12, 13 and references therein) suggests that multiple strategies for initiating XCI must exist. To explore XCI during early embryogenesis in other eutherian mammals, we analysed human (primate) and rabbit (lagomorph) embryos, the latter being closely related to rodents[14].

First we examined rabbit embryos *in vivo*, from the 4-cell stage to the blastocyst stage, using *Xist* RNA fluorescence *in situ* hybridization (FISH; Fig. 1 and Supplementary Fig. 1). Sex chromosome constitution was determined after RNA FISH, by DNA FISH using X- and Y-chromosome probes. Punctate *Xist* RNA signals could be detected from both X chromosomes in females and from the single X chromosome in males, from the 8-cell stage (Fig. 1b, d), which corresponds to major embryonic genome activation (EGA) in rabbits[15,16]. This contrasts with mice, where maternal *Xist* is never expressed[4,6,11] between major EGA (the 2-cell stage[17]) and the blastocyst stage. *Xist* RNA accumulation was first detected in female, but not male, embryos at the morula stage (62 hours post-coitum (h.p.c.); about 60–70 cells). However, the percentage of female blastomeres showing *Xist* expression was low at this stage (Fig. 1c, d). By the early blastocyst stage (96 h.p.c.; Fig. 2a), over 90% of trophectoderm cells had at least one *Xist* RNA domain (Fig. 2a, c). Surprisingly, a substantial proportion of blastomeres (26%) showed *Xist* RNA domains on both X chromosomes (Fig. 2a, c). The detection of double *Xist* RNA domains in rabbit blastocysts contrasts with mouse, where such a situation is not usually seen *in vivo*[4–6] except in polyploid trophectoderm cells or in a low percentage (<4%) of *in vitro* differentiating female embryonic stem cells[18]. This unusual pattern is transient, however, as it is no longer observed in trophectoderm cells of 120 h.p.c. blastocysts (Fig. 2b), where a single *Xist* RNA domain is seen, as in XX somatic cells (Supplementary Fig. 2). In male rabbit embryos, *Xist* RNA domains were rarely detected (Fig. 2c). Instead, *Xist* was progressively silenced after the 8-cell stage (Fig. 2b, c). Thus, *Xist* upregulation is sensitive to XX status in rabbits, as in mice during random XCI, but *Xist* monoallelic regulation differs between the two animals.

During mouse development, X-linked gene silencing follows *Xist* RNA accumulation[3–6]. However, the timing of XCI in other mammals is less clear. We investigated expression of two X-linked genes using RNA FISH in rabbit embryos: *Hprt1* (also known as *Hprt*), which is X-inactivated in somatic cells, and *Jarid1c* (*Kdm5c*), which escapes XCI (Supplementary Fig. 2b). Expression of both genes was biallelic in female embryos from the 4–16-cell stage (Fig. 1). Thus, X-linked genes are active from both Xp and Xm at EGA, as in mice[4,6].

Hprt1 silencing on the *Xist*-RNA-coated chromosome was first detectable at the 96-h.p.c. blastocyst stage and was complete in trophectoderm cells by the 120-h.p.c. stage (Fig. 2d), whereas *Jarid1c* still showed biallelic expression indicating escape from XCI, as in somatic cells. Importantly, the onset of *Hprt1* silencing coincided with the time window (96 h.p.c.) when both X chromosomes were sometimes found to be *Xist* RNA coated (Fig. 2a). Indeed, by comparing the proportion of cells showing no *Hprt1* primary transcript signal and two *Xist* RNA domains with those showing a single *Xist* RNA domain (46%, $n = 192$ versus 12%, $n = 250$), we conclude that *Hprt1* silencing has initiated on both X chromosomes in cells with two *Xist* RNA domains at the 96-h.p.c. stage (Fig. 2e). This situation was found to be rapidly reversed

[1]Mammalian Developmental Epigenetics Group, Institut Curie, CNRS UMR 3215, INSERM U934, Paris 75248, France. [2]Université Paris Diderot – Assistance Publique Hôpitaux de Paris – Service de Biologie de la Reproduction, Hôpital Bichat-Claude Bernard, 75018 Paris, France. [3]INRA UMR 1198 Biologie du Développement et de la Reproduction, F-78352 Jouy en Josas Cedex, France. [4]Laboratoire de Biologie de la reproduction CECOS, CHU de Dijon, Université de Bourgogne, 21079 Dijon, France. [5]Epigenetic Decisions and Reproduction in mammals, Institut Curie, CNRS UMR 3215, INSERM U934, 75248 Paris, France. [6]Unité Inserm 1016, Université Paris Descartes – Assistance Publique Hôpitaux de Paris – Service de Biologie de la Reproduction, Hôpital Cochin, 75014 Paris, France.
*These authors contributed equally to this work.

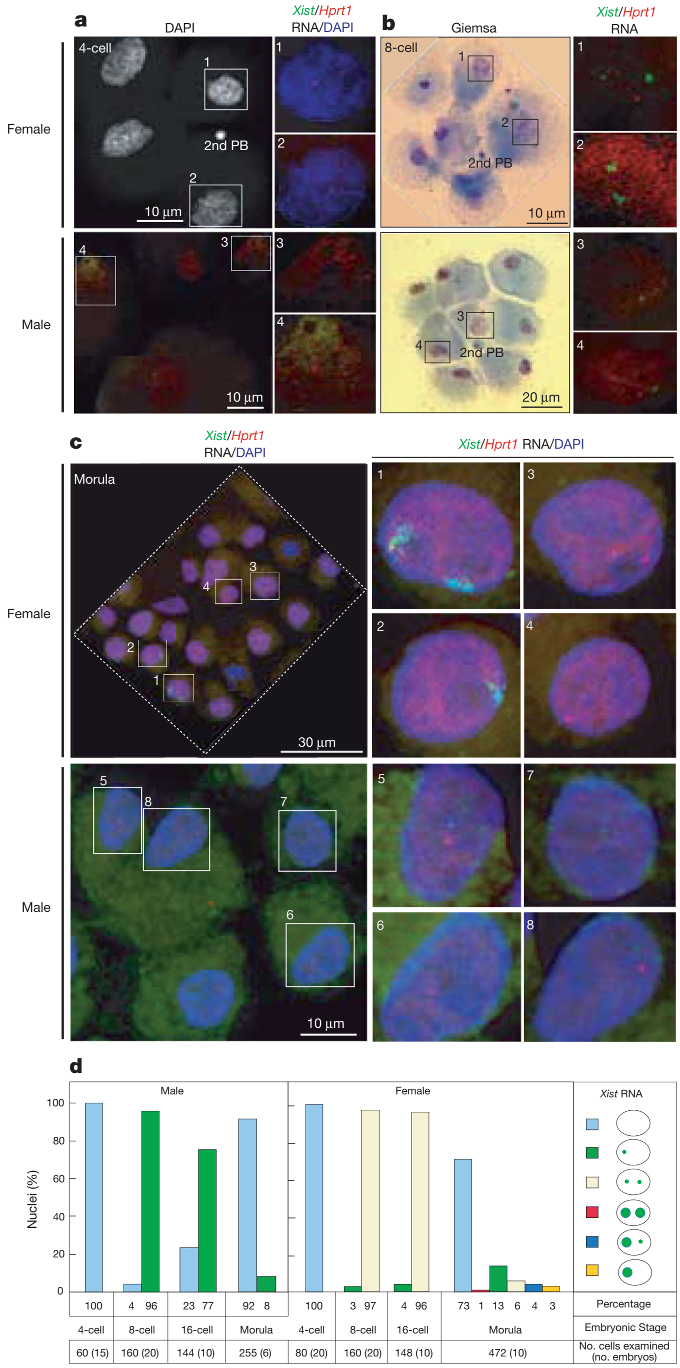


Figure 1 | *Xist* and *Hprt1* expression in early pre-implantation rabbit embryos. Examples of embryos after RNA FISH for *Xist* RNA (green) and X-linked *Hprt1* primary transcripts (red). An intact male or female embryo and several representative nuclei are shown. **a**, Representative 4-cell embryos: XX (nuclei 1 and 2) and XY (nuclei 3 and 4) blastomeres. DAPI, 4′,6-diamidino-2-phenylindole; PB, polar body. **b**, Representative 8-cell embryos: XX (nuclei 1 and 2) and XY (nuclei 3 and 4) blastomeres. **c**, Representative XX (nuclei 1–4) and XY (nuclei 5–6) morula-stage embryos. **d**, Percentage of cells showing different *Xist* expression patterns in the 4-cell to morula stages. The low percentage of *Xist* RNA detection in morulae is not a technical issue, as *Hprt1* primary transcript signals could be efficiently detected at this stage in most male and female blastomeres.

24 h later (at the 120-h.p.c. stage), as only one X chromosome showed *Xist* RNA coating and *Hprt1* silencing; the other expressed *Hprt1* in almost all blastomeres (Fig. 2b).

We also analysed histone H3 lysine 27 trimethylation (H3K27me3) enrichment, which marks the onset of XCI in mice[19,20] and is present on the inactive X chromosome in rabbit somatic cells (Supplementary Fig. 2c). In 96-h.p.c. blastocysts, where two *Xist*-RNA-coated X chromosomes could be detected (25%, $n = 325$), both of them showed H3K27me3 enrichment (Supplementary Fig. 3). At the 120-h.p.c. stage, where only one *Xist* RNA domain is detectable, two spots of H3K27me3 staining per nucleus could still sometimes be observed (8%, $n = 285$), one overlapping with *Xist* RNA and the other presumably corresponding to the second X chromosome (Supplementary Fig. 3).

Taken together, these results suggest that XCI initiates on both X chromosomes in a significant proportion of blastomeres in early rabbit blastocysts, but the situation is somehow resolved. As no differences in cell death rates could be detected between female and male 120-h.p.c. embryos (Supplementary Fig. 4), we conclude that cells initiating XCI on both X chromosomes do not necessarily die. They either reverse XCI on one X chromosome or proliferate more slowly and become out-competed by cells with one inactive X chromosome, which is a distinct possibility given the massive cell proliferation occurring between 96 h.p.c. (~100 cells) and 120 h.p.c. (>1,000 cells; Fig. 2f).

We also analysed the ICMs of rabbit blastocysts for XCI status. In mice, inactive Xp is progressively reactivated in pre-epiblast ICM cells[5,6] positive for Oct4 (Pou5f1) and Nanog[6]. Rabbit ICM cells were isolated from blastocysts by immunosurgery or manual dissection. In early blastocysts (96 h.p.c.), most putative ICM cells were negative for *Xist* RNA accumulation (Supplementary Fig. 5a, c), in contrast to trophectoderm cells, where *Xist* RNA coating of one or both X chromosomes was present in most cells (Fig. 2a). One day later (120 h.p.c.), however, the majority of putative (strongly Oct4-positive) epiblast cells displayed *Xist* RNA coating on one or both X chromosomes, as well as the onset of *Hprt1* gene silencing and H3K27me3 enrichment (Supplementary Fig. 5).

Thus, in rabbits, two waves of XCI with similar characteristics seem to occur, one in the trophectoderm and the second in the ICM. During both waves, the two X chromosomes seem to be affected initially in some cells, implying that there is no imprinting, unlike in the mouse. However, we could not analyse parent-specific X-linked gene expression owing to lack of polymorphisms in our rabbit colonies. Instead, we investigated rabbit parthenogenotes (two maternal genomes) for their XCI patterns. We found *Xist* and *Hprt1* expression patterns similar to those of wild-type female embryos (Supplementary Fig. 6), indicating that imprinting may not have a major role in rabbit XCI, but we cannot formally exclude that there is some parent-of-origin bias. In summary, we show that although rabbits and mice are phylogenetically close, they have taken very different evolutionary trajectories with respect to *Xist* regulation and the onset of XCI.

In human embryos, *XIST* expression is detected from the 4–8-cell stage and XCI has been proposed to follow kinetics similar to those in mouse[21], although it does not seem to be imprinted in extra-embryonic tissues[22,23]. To investigate this further, we examined *in vitro* conceived pre-implantation human embryos (Fig. 3; Methods and Supplementary Information). Human blastocysts showed either a single or double *XIST* RNA accumulation (67% cells, $n = 13$ embryos and 85% cells, $n = 15$ embryos, respectively; Fig. 3a), corresponding to male and female embryos, confirmed by X/Y DNA FISH (Fig. 3a). Similar *XIST* patterns were detected in morulae and blastocysts (days 5 and 6), whether derived by IVF (*in vitro* fertilization) or ICSI (intracytoplasmic sperm injection), and whatever their history (discarded or frozen-thawed embryos).

To assess XCI status, X-linked genes (*ATRX*, *FGD1* and *HUWE1*) were analysed by RNA FISH. In human somatic cells, these genes are silent on the inactive X chromosome (Supplementary Fig. 7). In female blastocysts, the majority of blastomeres showed two primary transcript foci for X-linked genes (*ATRX* and *FGD1*, Fig. 3a, e; *HUWE1*, data not shown), indicating that both X chromosomes are active. Furthermore, although small foci of H3K27me3 enrichment could be detected throughout the nucleus, no specific enrichment on the *XIST*-coated chromosomes was found (Fig. 3b).

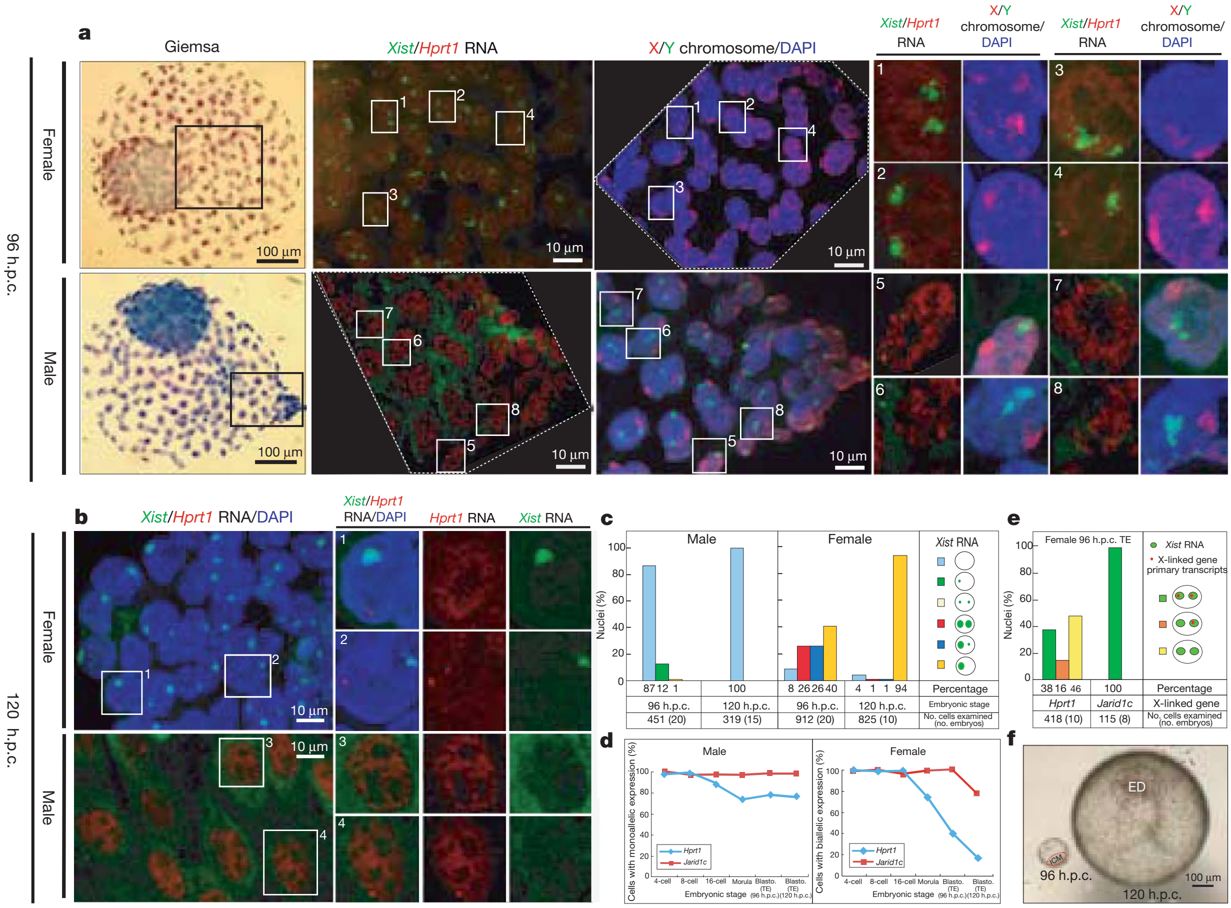


Figure 2 | *Xist* and *Hprt1* expression in early- (96-h.p.c.) and late-stage (120-h.p.c.) rabbit blastocysts. Examples of embryos (Giemsa-stained) analysed by RNA FISH to detect *Xist* (green) and *Hprt1* (red), followed by DNA FISH using X-chromosome (red) and Y-chromosome (green) probes. Intact male and female embryos, together with representative nuclei, are shown. **a**, Individual XX and XY early (96-h.p.c.) blastocysts: representative XX (nuclei 1–4) and XY (nuclei 5–8) cells. **b**, Late XX and XY blastocysts: representative XX (nuclei 1 and 2) and XY (nuclei 3 and 4) blastomeres. **c**, Percentage of cells showing different *Xist* expression patterns in rabbit embryos at early and late blastocyst stages. **d**, Average percentage of blastomeres showing primary transcripts for *Hprt1* or *Jarid1c* in female and male embryos from the 4-cell stage to the blastocyst stage ($n \geq 10$, each stage). TE, trophectoderm. **e**, Summary of X-linked gene expression patterns in blastomeres with double *Xist* RNA domains in early female blastocysts. **f**, Phase contrast images of living early (day-4, 96-h.p.c.) and late (day-5, 120-h.p.c.) rabbit blastocysts. ED, embryonic disc.

We also analysed a few even later human blastocysts (cultured to day 7). Male, day-7 blastocysts showed either no (66%, $n = 271$ cells, $n = 2$ embryos) or just a small punctate *XIST* RNA signal (Supplementary Fig. 8), suggesting that *XIST* expression was downregulated by this stage. A heterochromatic Y body[24] was also visible in these embryos. In the one female embryo obtained at day 7, the majority of cells still showed biallelic *FGD1* expression (83%, $n = 30$ cells) and biallelic *XIST* RNA signals, with no sign of a heterochromatic Barr body (Supplementary Fig. 8).

A previous study reported that XCI can be detected in early human embryos[21]. The discrepancy with our study may be due to differences in immunofluorescence/FISH techniques resulting in less efficient detection of biallelic RNA signals, for example, and/or to different culture conditions (see Supplementary Information for discussion). There is also some debate concerning XCI status in human embryonic stem cells. A recent study[25] reported that human embryonic stem cells cultured in 5% oxygen retain pre-XCI cells, whereas in atmospheric oxygen *XIST* upregulation and an inactive X chromosome are often found. However, this may be specific to the tissue culture conditions of human embryonic stem cells. We therefore analysed the ICM in both male and female human embryos. Putative ICM cells (strongly OCT4 positive) showed similar *XIST* RNA patterns to surrounding trophectoderm cells and no sign of an inactive X chromosome ($n = 4$ embryos; Fig. 3b). We conclude that in human pre-implantation embryos, under our conditions, *XIST* RNA upregulation occurs on all X chromosomes regardless of parental origin or sex but does not seem to result in chromosome-wide XCI even in late blastocysts. Although we cannot rule out that *in vitro* culture and atmospheric oxygen might have an impact on XCI, in our conditions the X seems to be in a pre-XCI state in human ICM cells, in agreement with ref. 25.

In conclusion, our findings reveal substantial diversity in the timing and regulation of XCI initiation between mammals (Supplementary Fig. 9a). First, *Xist* expression is not imprinted in early rabbit and human embryos, unlike in mice. The maternal *Xist* imprint found in mice may therefore be a recent event, which we have proposed[26] evolved to avoid premature maternal XCI given the early onset of EGA and *Xist* expression. Indeed, in mouse XpXp androgenotes, which lack an imprinted maternal Xm, *Xist* is upregulated biallelically initially, as in rabbits and humans[27]. However, rabbits and humans have later EGA and may not require *Xist* imprinting for early XCI control. Second, the choice of which X chromosome to inactivate seems to occur downstream of *Xist* upregulation in rabbits and humans, unlike in mice, where *Xist* is tightly monoallelically regulated during random XCI (Supplementary Fig. 9b). One difference could be that antisense transcription (*Tsix*), which is critical for random, monoallelic *Xist* expression in mice[28], is poorly conserved and may even be lacking in other mammals[29].

Another difference between human and rabbit or mouse embryos is that *XIST* upregulation occurs in males and females, suggesting that

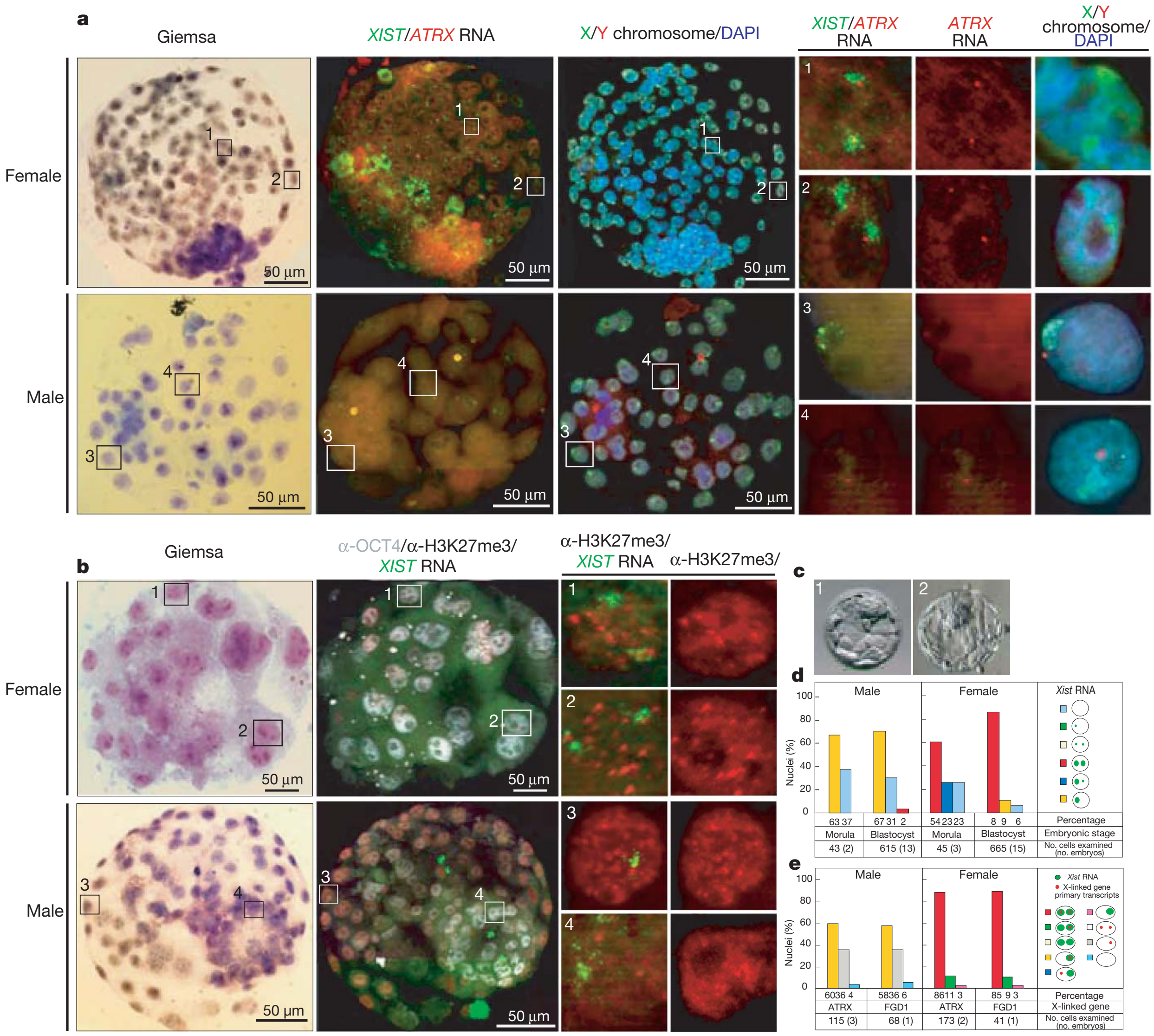


Figure 3 | *XIST* RNA, X-linked gene expression and H3K27me3 profiles in human blastocysts. **a**, Examples of individual female and male human embryos (Giemsa) analysed by RNA FISH for *XIST* RNA (green) and primary transcription from the X-linked *ATRX* gene (red), followed by DNA FISH for the X chromosome (red) and Y centromere (green). Representative XX (nuclei 1 and 2) and XY blastomeres (nuclei 3 and 4) are shown. Zones of intense green signal outside nuclei are background. **b**, Examples of individual female and male human embryos (Giemsa) analysed by immunolabelling with antibodies against H3K27me3 (red) and OCT4 (white), combined with *XIST* RNA FISH (green). For each sex, an intact embryo (Giemsa, immunofluorescence/RNA FISH) together with several enlarged nuclei are shown: representative XX (nuclei 1 and 2) and XY blastomeres (nuclei 3 and 4). **c**, Phase contrast images of an expanding blastocyst (left) and a fully expanded blastocyst (right). **d**, Percentage of cells showing different *XIST* expression patterns in morula- to blastocyst-stage embryos. **e**, Percentage of female and male blastocyst cells showing different X-linked gene expression patterns in blastomeres with *XIST* RNA domains.

XIST activators are not XX-specific, at least initially (Supplementary Fig. 9a). However, *XIST* upregulation is dissociated from the immediate onset of chromosome-wide XCI, possibly as a result of incomplete *XIST* RNA *cis*-coating. Indeed, *XIST* RNA signals were quite diffuse in human embryos (Fig. 3) in comparison with somatic cells (Supplementary Fig. 7). Importantly, the lack of chromosome-wide XCI in the trophectoderm of human blastocysts implies that dosage compensation may not be critical at this stage.

Finally, we also reveal differences between mammals in XCI status in the ICM. In humans the two X chromosomes seem to be active in the ICM (even though *XIST* is expressed), whereas in rabbits the two X chromosomes are initially active and then become inactive. This is the reverse situation to mice, where Xp is initially inactive and then becomes reactivated in the ICM[6]. These differences may explain the diverse XCI patterns reported in embryonic stem cells of non-murine mammals and point to differences in *Xist* control via pluripotency factors[30]. The diversity of XCI initiation strategies we have uncovered here probably reflects the fact that developmental processes are constantly changing during evolution and that the regulation of processes such as XCI have to display substantial plasticity to accommodate these changes.

METHODS SUMMARY

Human and rabbit embryos were manipulated under strict ethical guidelines and analysed using published protocols for immunostaining, RNA and DNA FISH[6,26]. Human embryos obtained following IVF or ICSI were cultured in 5% CO_2 and atmospheric O_2, fixed and used for immunofluorescence or RNA FISH, as previously described[6]. Rabbit embryos were obtained from superovulated New Zealand White rabbit females mated with males. *In vivo* rabbit embryos were recovered from oviducts or uterus flushed with PBS at appropriate times post-coitum. The mucin coat and zona pellucida were removed before fixation for immunofluorescence/FISH[6]. Rabbit ICMs (96 h.p.c.) were isolated by immunosurgery; embryonic discs (120 h.p.c.) were dissected out with tungsten needles.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 August 2010; accepted 27 January 2011.

Published online 6 April 2011.

1. Lyon, M. F. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190,** 372–373 (1961).
2. Sharman, G. B. Late DNA replication in the paternally derived X-chromosome of female kangaroos. *Nature* **230,** 231–232 (1971).
3. Huynh, K. D. & Lee, J. T. Inheritance of a pre-inactivated paternal chromosome in early mouse embryos. *Nature* **426,** 857–862 (2003).
4. Patrat, C. *et al.* Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl Acad. Sci. USA* **106,** 5198–5203 (2009).
5. Mak, W. *et al.* Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303,** 666–669 (2004).
6. Okamoto, I., Otte, A. P., Allis, C. D., Reinberg, D. & Heard, E. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303,** 644–649 (2004).
7. Takagi, N. & Sasaki, M. Preferential inactivation of the paternally derived X-chromosome in the extraembryonic membranes of the mouse. *Nature* **256,** 640–642 (1975).
8. Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. & Brockdorff, N. Requirement for *Xist* in X chromosome inactivation. *Nature* **379,** 131–137 (1996).
9. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. *Xist*-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11,** 156–166 (1997).
10. Kay, G. F., Barton, S. C., Surani, M. A. & Rastan, S. Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell* **77,** 639–650 (1994).
11. Navarro, P. & Avner, P. An embryonic story: analysis of the gene regulative network controlling Xist expression in mouse embryonic stem cells. *Bioessays* **32,** 581–588 (2010).
12. Duret, L., Chureau, C., Samain, S., Weissenbach, J. & Avner, P. The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science* **312,** 1653–1655 (2006).
13. Okamoto, I. & Heard, E. Lessons from comparative analysis of X-chromosome inactivation in mammals. *Chromosome Res.* **17,** 659–669 (2009).
14. Hajjoubi, S. *et al.* Ruminants genome no longer contains Whey Acidic Protein gene but only a pseudogene. *Gene* **370,** 104–112 (2006).
15. Manes, C. The participation of the embryonic genome during early cleavage in the rabbit. *Dev. Biol.* **32,** 453–459 (1973).
16. Christians, E., Rao, V. H. & Renard, J. P. Sequential acquisition of transcriptional control during early embryonic development in the rabbit. *Dev. Biol.* **164,** 160–172 (1994).
17. Flach, G. *et al.* The transition from maternal to embryonic control in the 2-cell mouse embryo. *EMBO J.* **1,** 681–686 (1982).
18. Monkhorst, K., Jonkers, I., Rentmeester, E., Grosveld, F. & Gribnau, J. X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. *Cell* **132,** 410–421 (2008).
19. Plath, K. *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300,** 131–135 (2003).
20. Silva, J. *et al.* Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* **4,** 481–495 (2003).
21. van den Berg, I. M. *et al.* X chromosome inactivation is initiated in human preimplantation embryos. *Am. J. Hum. Genet.* **84,** 771–779 (2009).
22. Migeon, B. X-chromosome inactivation: theme and variations. *Cytogenet. Genome Res.* **99,** 8–16 (2002).
23. Moreira de Mello, J. C. *et al.* Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. *PLoS ONE* **5,** e10947 (2010).
24. Barlow, P. & Vosal, C. G. The Y chromosome in human spermatozoa. *Nature* **226,** 961–962 (1970).
25. Lenger, C. J. *et al.* Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* **141,** 1–12 (2010).
26. Okamoto, I. *et al.* Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438,** 369–373 (2005).
27. Okamoto, I., Tan, S. & Takagi, N. X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation. *Development* **127,** 4137–4145 (2000).
28. Lee, J. T. Regulation of X-chromosome counting by *Tsix* and *Xite* sequences. *Science* **309,** 768–771 (2005).
29. Chang, S. C. & Brown, C. J. Identification of regulatory elements flanking human XIST reveals species differences. *BMC Mol. Biol.* **11,** 20 (2010).
30. Navarro, P. *et al.* Molecular coupling of *Xist* regulation and pluripotency. *Science* **231,** 1693–1695 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank members of the Heard lab for advice and discussions, V. Colot for comments on the manuscript, and P. Jouannet and Hôpital Cochin's IVF unit for support and advice on human embryo experiments. We are indebted to the Unité Commune d'Expérimentation Animale in charge of the rabbit colonies, C. Archilla for technical help in sexing rabbit embryos, C. Rogel-Gaillard for providing rabbit bacterial artificial chromosomes before publication, and the UMR3215 PICT-IBiSA imaging facility. Funding is from the FRM (Equipe FRM), the ANR and the ERC (E.H.); the Agence de Biomedecine (C.P. and P.F.); and the INRA PHASE department (ACI 2007) (V.D. and J.-P.R.).

Author Contributions E.H., I.O., J.-P.R., V.D. and C.P. designed the experiments. I.O. performed immunofluorescence and FISH experiments on both rabbit and human embryos; C.P. and P.D. performed some of the FISH experiments on human embryos. C.P. obtained the license to work with human embryos and interviewed couples to obtain consent for use of discarded embryos. C.P. and P.F. performed the human IVF and ICSI experiments and manipulated the human embryos. J.-P.W. directed the laboratory in which the human embryos were made available. D.T. screened rabbit bacterial artificial chromosome libraries, provided rabbit probes and conducted comparative sequence analysis. N.P., N.D. and V.D. obtained and manipulated rabbit embryos. J.-P.R. directed the laboratory in which the rabbit embryos were obtained and manipulated. I.O., C.P., D.T., V.D. and E.H. analysed the data. I.O. and E.H. wrote the manuscript together with input from V.D. and C.P.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.H. (edith.heard@curie.fr) or V.D. (veronique.duranthon@jouy.inra.fr).

METHODS

Recovery of pre-implantation rabbit embryos. Experiments were performed in accordance with the International Guiding Principles for Biomedical Research involving animals, as promulgated by the Society for the Study of Reproduction and with the European Convention on Animal Experimentation. Researchers involved in direct work with the animals possessed an animal experimentation licence delivered by the French veterinary services. New Zealand White rabbit females (INRA line 1077) were superovulated and mated with normal males. Superovulation was performed by five subcutaneous injections of pFSH (Stimufol, Merial Lyon) for three days before mating: two 5-μg injections on day 1 at 12-h intervals, two 10-μg injections on day 2 at 12-h intervals and one 5-μg injection on day 3, followed 12 h later by an intravenous injection of 30 IU human chorionic gonadotrophin (Chorulon, Intervet) at mating time. *In vivo*, four-cell, eight-cell, sixteen-cell and morula-stage embryos were recovered from oviducts and uterus flushed with PBS at 30, 38, 46, 62–72 h.p.c., respectively. Uterus flushing was at 96 and 120 h.p.c. for the recovery of early and late blastocysts, respectively. Mucin coat was removed by incubation of the embryos in a pre-warmed pronase solution (0.5% Protease, Sigma) at 37 °C for 1 to 6 min depending on its thickness and on the stage of the embryo. Embryos were then immediately rinsed twice in M199 HEPES with 10% FCS. In the second wash, the zona pellucida was removed by mechanical pipetting. For immunofluorescence and RNA FISH, embryonic discs were dissected out from the late (120-h.p.c.) blastocysts with tungsten needles.

ICM isolation from early (96-h.p.c.) rabbit blastocyst. Mucin coat and zona pellucida were removed by pronase treatment of blastocysts. ICMs were then isolated by immunosurgery. Briefly, embryos were incubated in anti-rabbit goat serum (Sigma R-5131) for one hour at 37 °C. They were then rinsed twice in PBS and incubated in guinea pig complement (Sigma S-1639) for one to two minutes. Blastocysts were then rinsed in PBS and vacuolated trophectoderm cells were mechanically removed by pipetting with a small adjusted pipette (70–90-μm diameter).

Parthenogenetic activation of rabbit embryos. *In vitro* activation was performed on morphologically normal, fresh MII oocytes (14 h.p.c.), as previously described[31,32] for nuclear transfer experiments. Briefly, oocytes were washed three times in the TCM199, HEPES plus 10% FCS and placed in a mannitol solution containing 0.3 M mannitol, 100 μM $MgCl_2$ and 100 μM $CaCl_2$ in sterile, filtered water. After 30 s in the mannitol solution, oocytes were 'electroactivated' with a succession of three d.c. pulses of 3.2 kV cm^{-1} for 20 μs each, using a BTX stimulator (Biotechnologies & Experimental Research). Then oocytes were incubated for 1 h in TCM199 plus 10% FCS at 38.5 °C and 5% CO_2 in a humidified atmosphere. Then they received, in the mannitol solution, a second set of three d.c. pulses and were incubated for 1 h in TCM199 plus 10% FCS supplemented with 2 mM 6-DMAP (kinase inhibitor) and 5 μg μl^{-1} cycloheximid (protein synthesis inhibitor) to finalize activation. Lastly, they were washed three times in the TCM199 plus 10% FCS to remove all traces of chemical agents, and were rinsed for 30 min in TCM199 plus 10% FCS at 38.5 °C and 5% CO_2. They were then incubated in microdrops of B2 plus 2.5% FCS under mineral oil at 38.5 °C and 5% CO_2.

Recovery of human pre-implantation embryos. All attempts using assisted reproductive techniques were performed at Cochin-Saint Vincent Hospital (Paris, France). Ovarian stimulation, oocyte retrieval, sperm processing and IVF and/or ICSI were performed as previously described[33]. The choice of ICSI or conventional IVF as fertilization method was dependent on semen-sample characteristics and couple's history. Briefly, oocytes retrieval was performed under vaginal ultrasound guidance, 36 h after administration of human chorionic gonadotrophin. Motile spermatozoa were selected through a two-step 90–45% density gradient (Puresperm, JCD) or after sperm washing with FertiCult medium (JCD). On the day of oocyte retrieval, sperm-inseminated or -microinjected oocytes were cultured in 30 μl culture medium (IVF or ISM1, Origio) under oil, at 37 °C, 5% CO_2 and humidified atmosphere. Fertilization was assessed 18 h after insemination or injection and normal fertilization was defined as the presence of two distinct pronuclei and a second polar body. Embryo quality was evaluated on day 2 or day 3 according to the number of blastomeres, the degree of fragmentation and the presence or not of multinuclear blastomeres. Embryos were either transferred or cryopreserved on day 2 or day 3, depending on the embryo quality evaluation. Embryos that were neither transferred nor cryopreserved, given poor embryo morphology, were discarded.

Collection of human embryos. Research on human embryos was authorized by the French Biomedicine Agency (RE07-011R). In the present study, only diploid embryos that were either discarded or cryopreserved were used. In both cases, written consent was obtained from the couples allowing these discarded or cryopreserved embryos to be used for the research. Embryos (after thawing (Embryo Thawing Pack, Origio), for day-2–3 cryopreserved embryos) were placed in 30 μl of fresh, equilibrated sequential culture medium (ISM1 culture medium from day 2 to day 3 and ISM2 from day 3, Medicult) at 37 °C, 5% CO_2, under oil and humidified atmosphere, for further culture. Blastocysts were obtained at day 5 or day 6 post-insemination or -injection and classified according to Gardner's classification[34], taking into account the global morphology and ICM and trophectoderm aspect from mid blastocyst.

Rabbit probes. All rabbit bacterial artificial chromosomes (BACs) were isolated from a rabbit library maintained at the UR1313, Unité de Génétique Animale et Biologie Intégrative, Institut National de la Recherché Agronomique (INRA). *Hprt1* and *Jarid1c* BACs had been previously published[35,36]. *Xist*-containing BACs (573D5, 614B8 and 717B2) were isolated by a PCR-based screening protocol (C. Rogel-Gaillard, unpublished results). BACs were verified for their location and single-copy nature by DNA FISH on metaphase spreads. For *Xist* RNA FISH, shorter probes were derived from the BACs. First, the human *XIST* gene sequence was compared with discontinuous megaBLAST to rabbit whole-genome shotgun sequences from the Trace Archives of the NCBI site to identify homologous rabbit *Xist* sequences. Second, these genomic sequences were extracted and then assembled in six large sequences using the CAP3 assembly program[37]. These six large sequences encompass the promoter and beginning of exon 1 (4,090 base pairs (bp)), the middle of exon 1 (6,255 bp), the end of exon 1 and part of intron 1 (3,250 bp), from the end of intron 4 to more than 3 kb into exon 6 (5,792 bp), the middle of exon 6 (1,300 bp) and the end of exon 6 (2,067 bp), respectively. Finally, primers were designed to amplify large parts of the rabbit *Xist* gene from the BACs. For RNA FISH, a mixture of these three probes obtained from long-range PCR products (Roche) were amplified using the 8F/17R, 16F/3R and 6F/6R primers pairs, respectively (8F, 5′-TATAGGAATGCTGACCACTGA-3′; 17R, 5′-GTTTTTTCAGTCAAGGGAGAC-3′; 16F, 5′-TCATGCTCCTGAATCTTCTTG-3′; 3R, 5′-GCCGCACAAGGAAGAAAAAAG-3′; 6F, 5′-TGTGTATTTCTCTTTGTCTTT-3′; 6R, 5′-GTTCTCAAAGCAGTACATTTTTATTTTACAGTACACAAC 3′). The three PCR products encompass sequences from the middle of exon 1 to the end of intron 1 (~8 kb), the beginning of intron 1 to the beginning of exon 6 (~10 kb), and all of exon 6 (~7.3 kb) (Supplementary Fig. 1).

Human probes. BAC probes (Supplementary Table 2) were used following verification of their location and single-copy nature by DNA FISH on female human metaphases. Two DNA probes detecting *XIST* RNA were used: (i) a 10-kb fragment corresponding to *XIST* exon 1 (gift from C. Brown, Department of Medical Genetics, University of British Columbia) and (ii) a VI.34 Flp-In T-Rex *XIST* CDNA DH10B construct[38] (gift from C. Brown).

Immunofluorescence, RNA FISH and DNA FISH. Immunofluorescence and RNA FISH were carried out as described previously[6,26]. DNA FISH was performed following RNA FISH. The coverslips carrying the embryos were then recovered in ×2 SSC and incubated in RNase A (100 μg ml^{-1}) in ×2 SSC at 37 °C for 1 h, before a brief rinse in ×2 SSC, dehydration through an ethanol series and air drying. The coverslips were treated with 0.03% pepsin in 0.01 N HCl for 10–30 min at 37 °C to enhance the penetration of probes. After washing in PBS, they were dehydrated through an ethanol series. The DNA was then denatured in 70% formamide, ×2 SSC (pH 7.4) for 10 min at 80 °C and dehydrated again through an ice-cold ethanol series. Hybridization and revelation of the chromosome paint probes were performed according to manufacturer's instructions (Cambio). Cover slips were counterstained with DAPI (1 μg ml^{-1}), mounted and viewed under the fluorescence microscope.

TUNEL assay and sexing of rabbit embryos. *In vivo* developed blastocysts were recovered at 120 h.p.c. They were dissected into two halves. One half of each embryo was used for sexing by nested PCR against the *Sry* gene. Briefly, frozen hemi-embryos were lysed in 10 μl of TE buffer containing 0.01 mg ml^{-1} proteinase K and incubated in this buffer for 2 h at 56 °C. Proteinase K was then inactivated by heating at 94 °C for 10 min. PCR was then directly performed in a 50-μl final volume using Takara Taq DNA polymerase (Lonza) for 30 cycles using external primers (*Sry* F2 5′-gttcggagcactgtacagcg-3′ and *Sry* R2 5′-gcgttcatgggtcgcttgac-3′). A quantity (3 μl) of the first PCR reaction was then re-amplified in a 50-μl PCR reaction for 35 cycles using the internal primers (*Sry* F1 5′-tgcttacacaccagccaaaca-3′ and *Sry* R1 5′-ttcctggccgctcactttac-3′) (PCR cycles: 94 °C for 3 min, then 94 °C for 30 s, 58 °C for 30 s and 72 °C for 1 min). PCR products (20 μl) were analysed by electrophoresis on a 2% agarose gel. The other half of each embryo was underwent TUNEL assay using the DeadEnd Fluorometric TUNEL System (Promega). Briefly, hemi-embryos were fixed in paraformaldehyde (4%) overnight at 4 °C. After rinsing in PBS containing 0.2% BSA, they were incubated in 10 μg ml^{-1} Proteinase K for 2 min. After rinsing, they were fixed again in 4% paraformaldehyde and 0.2% glutaraldehyde for 15 min at room temperature (22 °C). They were then rinsed twice in PBS containing 0.2% BSA. Positive control embryos were treated by RQ1 RNase-free DNase for 10 min as recommended by the manufacturer. TUNEL reaction was then performed on all the embryos as recommended by the manufacturer. The embryos were then incubated for 10 min at room temperature in

$10\,\mu g\,ml^{-1}$ Hoechst (Sigma). They were rinsed in 0.2% BSA in PBS and mounted in Citifluor (Biovalley) before imaging.

Fluorescence microscopy. A 200M Axiovert (Zeiss) fluorescence microscope equipped with an ApoTome system was used for image acquisition and the generation of optical sections in three dimensions. Sequential *z*-axis images were collected in 0.3-μm steps. At the blastocyst stage, when possible, we distinguished cells corresponding to the trophectoderm or the ICM according to their morphology. Results were expressed as mean ± s.e.m.

31. Chesné, P. *et al.* Cloned rabbits produced by nuclear transfer from adult somatic cells. *Nature Biotechnol.* **20,** 366–369 (2002).
32. Challah-Jacques, M., Chesne, P. & Renard, J. P. Production of cloned rabbits by somatic nuclear transfer. *Cloning Stem Cells* **5,** 295–299 (2003).
33. Fauque, P. *et al.* Cumulative results including obstetrical and neonatal outcome of fresh and frozen-thawed cycles in elective single versus double fresh embryo transfers. *Fertil. Steril.* **94,** 927–935 (2010).
34. Gardner, D. K. & Schoolcraft, W. B. in *Towards Reproductive Certainty: Infertility and Genetics Beyond 1999* (eds Jansen, R. & Mortimer, D.) 378–388 (Parthenon, 1999).
35. Hayes, H. *et al.* Establishment of R-banded rabbit karyotype nomenclature by FISH localization of 23 chromosome-specific genes on both G- and R-banded chromosomes. *Cytogenet. Genome Res.* **98,** 199–205 (2002).
36. Chantry-Darmon, C. *et al.* 133 new gene localization on the rabbit cytogenetic map. *Cytogenet. Genome Res.* **103,** 192–201 (2003).
37. Huang, X. & Madan, A. CAP3: A DNA sequence assembly program. *Genome Res.* **9,** 868–877 (1999).
38. Chow, J. C. *et al.* Inducible XIST-dependent X-chromosome inactivation in human somatic cells is reversible. *Proc. Natl Acad. Sci. USA* **104,** 10104–10109 (2007).

doi:10.1038/nature09869

Genome-wide analysis reveals novel molecular features of mouse recombination hotspots

Fatima Smagulova[1]*, Ivan V. Gregoretti[2]*, Kevin Brick[2], Pavel Khil[2], R. Daniel Camerini-Otero[2] & Galina V. Petukhova[1]

Meiotic recombination predominantly occurs at discrete genomic loci called recombination hotspots, but the features defining these areas are still largely unknown (reviewed in refs 1–5). To allow a comprehensive analysis of hotspot-associated DNA and chromatin characteristics, we developed a direct molecular approach for mapping meiotic DNA double-strand breaks that initiate recombination. Here we present the genome-wide distribution of recombination initiation sites in the mouse genome. Hotspot centres are mapped with approximately 200-nucleotide precision, which allows analysis of the fine structural details of the preferred recombination sites. We determine that hotspots share a centrally distributed consensus motif, possess a nucleotide skew that changes polarity at the centres of hotspots and have an intrinsic preference to be occupied by a nucleosome. Furthermore, we find that the vast majority of recombination initiation sites in mouse males are associated with testis-specific trimethylation of lysine 4 on histone H3 that is distinct from histone H3 lysine 4 trimethylation marks associated with transcription. The recombination map presented here has been derived from a homogeneous mouse population with a defined genetic background and therefore lends itself to extensive future experimental exploration. We note that the mapping technique developed here does not depend on the availability of genetic markers and hence can be easily adapted to other species with complex genomes. Our findings uncover several fundamental features of mammalian recombination hotspots and underline the power of the new recombination map for future studies of genetic recombination, genome stability and evolution.

The vast majority of homologous recombination takes place in recombination hotspots—discrete regions of the genome with a recombination frequency significantly above the frequency in adjacent areas. Extensive studies of several individual hotspots in mammals have greatly advanced our understanding of hotspot biology (reviewed in refs 1–6), but some critical features may be missed without examining the full ensemble of hotspots in a genome. Recently, remarkable progress has been made that culminated in the identification of more than 30,000 recombination hotspots in human[7–11]. Nevertheless, the relatively low resolution of the human map and the high variability of the recombination pattern between individuals still hinders fine structural analysis of hotspots. At the same time, it is clear that the primary DNA structure per se is a poor predictor of recombination activity and that hotspot position depends on additional factors, including epigenetic marks and most probably others[1–4]. Therefore, defining the complete set of variables that ultimately determine hotspot sites in mammals would be greatly facilitated if a hotspot map of a genetically homogeneous and malleable organism were available. We therefore embarked on generating a high-resolution physical map of recombination hotspots in the mouse using a direct molecular approach for identification of recombination initiation sites.

Meiotic recombination is initiated by the introduction of DNA double-stranded breaks (DSBs) by the protein SPO11, followed by resection of the ends to produce long, single-stranded overhangs[12]. Proteins RAD51 and DMC1 form nucleoprotein filaments at the ends of the breaks and search for a homologous chromosome that is used for repair[12]. We used anti-DMC1 antibodies to localize recombination initiation sites in the male mouse genome by chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq). To enrich for DSB-stage spermatocytes, we used $Hop2^{-/-}$ mice (*Hop2* also known as *Psmc3ip* or *Tbpip*) that lack the cells of later spermatogenic stages owing to meiotic arrest after DSB formation[13] (Supplementary Text). Hotspots identified by our approach in wild-type and in $Hop2^{-/-}$ mice correlate extensively (Supplementary Fig. 1), but four times more hotspots can be identified using $Hop2^{-/-}$ mice, owing to a higher signal-to-noise ratio.

Consistent with co-localization of RAD51 and DMC1 to DSB sites[12], we found that tag coverage in anti-RAD51 ChIP-Seq was highly correlated with that in anti-DMC1 ChIP-Seq (Fig. 1). Furthermore, the mapping data were highly reproducible between biological replicates ($R = 0.71$–0.97 (correlation coefficient); Fig. 1c). We were able to identify 9,874 recombination hotspots ($P = 10^{-4}$, FDR = 6.7% (false-discovery rate); Supplementary Data and Supplementary Figs 2 and 3), although the number of hotspots could be higher when less restrictive parameters are used (Supplementary Fig. 4). Confirmation of several identified hotspots was carried out by two approaches including direct physical detection of DSBs as previously described[14] (Supplementary Fig. 5). Furthermore, we found that the correlation of our DSB map with available genetic maps[15,16] is almost as high as the correlation between the genetic maps themselves (Fig. 1d, Supplementary Fig. 6 and Supplementary Text). The correlation of these genetic maps with the DSB hotspot distribution from this study is an important validation of our hotspot mapping approach. Additional supporting evidence comes from our finding that the 'hottest' cluster of DSB hotspots in the mouse genome is located in the PAR—the only homologous region between the X and Y chromosomes (Fig. 1e and Supplementary Fig. 3). Despite the very short length of the PAR, each spermatocyte undergoes an obligatory crossover in this area[17]. Estimates from our analysis show that this cluster is probably sufficient to ensure that there is at least one DSB in the PAR of every spermatocyte (Supplementary Text).

The centres of the hotspots in our map are defined with an approximate precision of 200 nucleotides, which is at least an order of magnitude higher than that of other available recombination maps in multicellular organisms (Supplementary Methods). More than 250 of the identified hotspots are hotter than one of the hottest previously known hotspots, *H2Eα* (Fig. 2a). On the basis of the estimated maximum recombination frequency of the *H2Eα* hotspot, 2 cM (ref. 18), the strength of the hottest hotspot in the mouse might be as high as 6 cM. Most of the hotspots lie 60–330 kb apart, with only a few recombination deserts more than 3 Mb in length (Supplementary Fig. 7). The average strengths of individual hotspots on different autosomes are similar, and on a chromosome scale the slight variation in hotspot density between autosomes does not correlate with the density of genes, chromosomal GC content, abundance of DNA repeats or chromosome length (not shown).

[1]Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814, USA. [2]National Institute of Diabetes, Digestive and Kidney Diseases, NIH, Bethesda, Maryland 20892, USA.
*These authors contributed equally to this work.

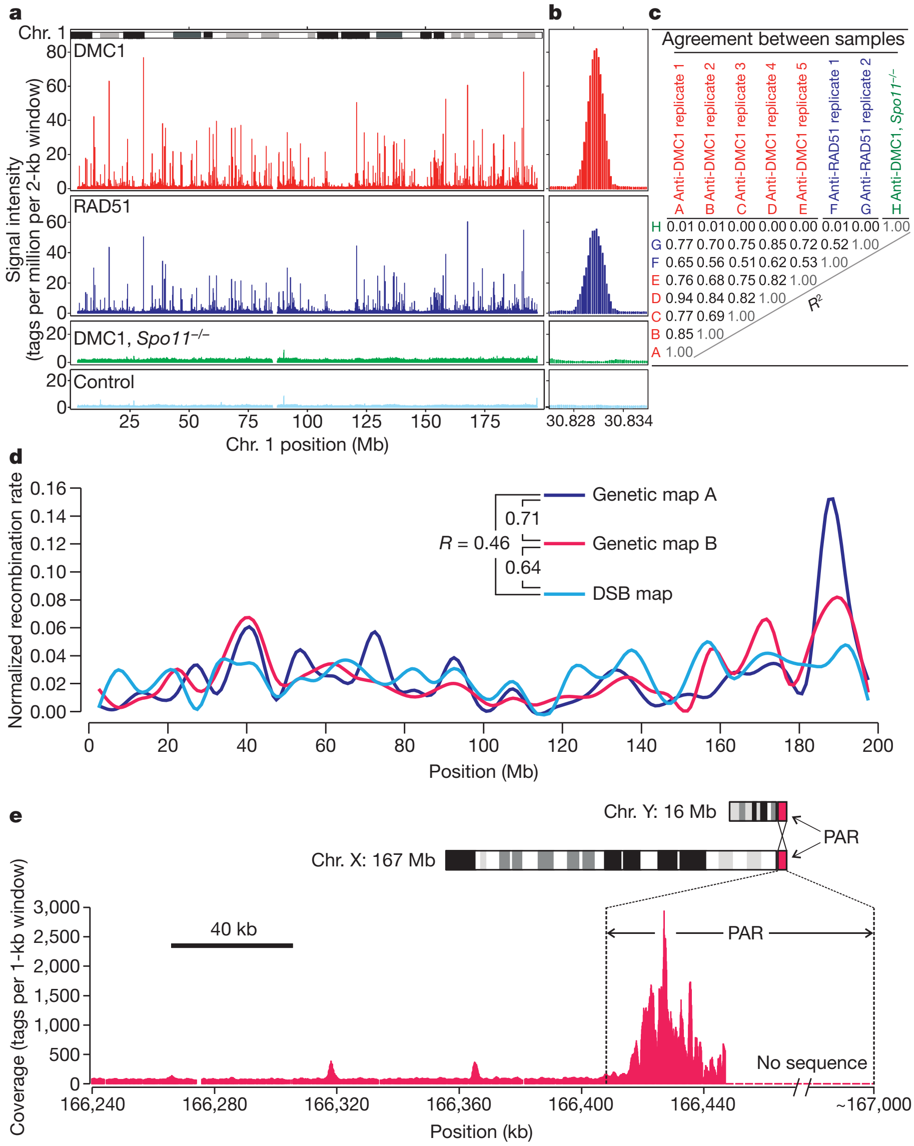


Figure 1 | DSB hotspots in the mouse genome. a, ChIP-Seq tag density profiles. DMC1: anti-DMC1 ChIP; RAD51: anti-RAD51 ChIP; DMC1, *Spo11*$^{-/-}$: anti-DMC1 ChIP from *Spo11*$^{-/-}$ mice that do not form DSBs; Control: IgG ChIP and input DNA pool. **b**, Close-up of a representative hotspot. **c**, Agreement between ChIP-Seq samples (correlations in 2-kilobase (kb) bins across genome). **d**, Correlation between the DSB hotspot map and the published genetic maps A[15] and B[16] for chromosome 1. The DSB map is generated from hotspot strengths (Supplementary Methods). All maps are generated in 5-megabase (Mb) windows and normalized by area of the map. **e**, The pseudoautosomal region (PAR) contains a large cluster of overlapping hotspots. DMC1 ChIP-Seq tag coverage is shown (smoothing window, 1 kb; step size, 100 bp).

We found that mouse hotspots have a tendency to overlap genes (Fig. 2b), although only the hottest 40% of the hotspots contribute significantly to this correlation (Supplementary Fig. 8). We next examined the association of DSB hotspots with different classes of DNA repeats and other genomic features. Significant correlations were found with GC content, short and long interspersed nuclear elements, long terminal repeats and other repeats, in agreement with those found in previous work[1,2,4,5] (Supplementary Table 1 and Supplementary Text).

In searching for additional features that might define hotspot locations, we analysed the nucleotide composition of the hotspot regions, as characteristic nucleotide skews have been found at some functional genomic elements including replication origins and transcription start sites[19]. Examination of either single strand of the double-stranded DNA in the 5′-to-3′ direction reveals that the sequence 5′ to the centre of a hotspot is enriched in purines but that the polarity of the bias changes in the middle of the hotspot, such that the sequence 3′ to the hotspot centre is more rich in pyrimidines (Fig. 2c and Supplementary Fig. 9). Replication- and transcription-related skews have been attributed to mutational biases acting asymmetrically on complementary DNA strands. It is conceivable that the skew detected at hotspots is the result of mutational asymmetry as well (Supplementary Text and Supplementary Fig. 9c). It is also possible that the skew represents some unknown functional feature of the genome that favours DSB formation. Importantly, we were able to detect the same signature of nucleotide usage in human hotspots (Supplementary Fig. 9d), suggesting that the purine–pyrimidine skew is an intrinsic property of recombination hotspots in mammals. We also noticed a slight increase in the overall GC content in the middle of hotspots (not shown), which might indicate the presence of a gene conversion bias[20].

The protein PRDM9 is a meiosis-specific methyltransferase responsible for trimethylation of lysine 4 on histone H3[21] (H3K4me3). In addition to its well-known role in transcription[22], this histone modification is associated with increased recombination activity in yeast and in mouse[23,24]. PRDM9 is highly polymorphic in its multi-Zn^{2+}-finger DNA-binding domain, and recent studies have implicated PRDM9 in determining the positions of recombination hotspots through the

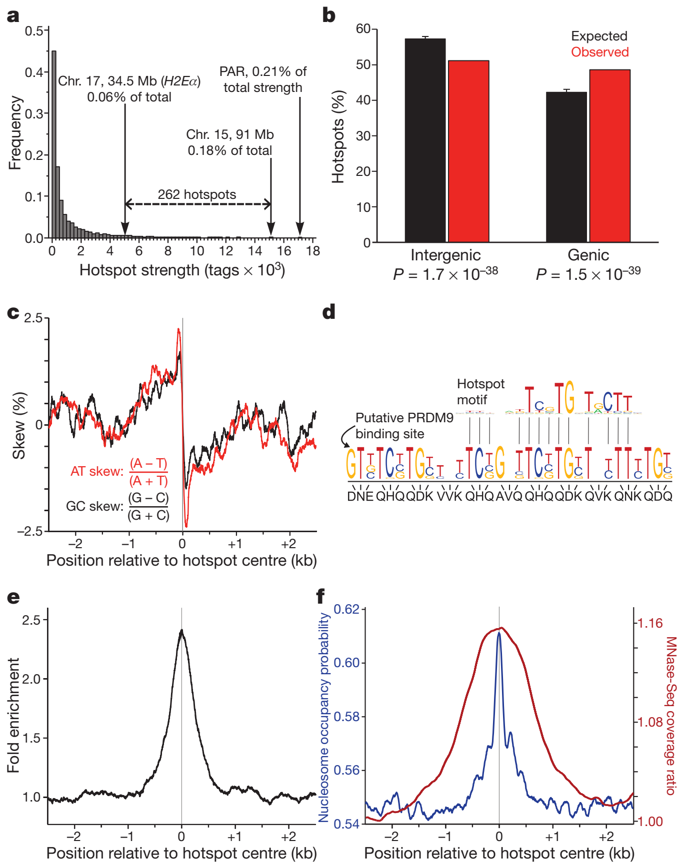


Figure 2 | Characteristics of mouse DSB hotspots. **a**, The *H2Eα* hotspot and the PAR hotspot cluster are among the strongest in the mouse genome. The strength of the hottest individual hotspot identified in this study (chromosome 15, ~91 Mb) is also indicated. **b**, Mouse DSB hotspots are significantly enriched in genes (one-sided binomial tests). Genic regions are defined from start to stop codons including introns. Error bars, 5th to 95th percentiles of the expected value distributions ($n = 10{,}000$ iterations). **c**, A purine–pyrimidine skew is apparent at DSB hotspots. Skew is calculated in 100-bp windows with a step size of 1 bp. **d**, The consensus hotspot motif is similar to the predicted binding site of PRDM9. **e**, The consensus motif is present in the centres of DSB hotspots. The distribution of hits to the consensus motif is shown in the 5-kb regions around hotspots (window size, 200 bp; step size, 1 bp). **f**. Both predicted (blue) and experimentally determined (red) nucleosome occupancy profiles peak at the centres of DSB hotspots. MNase-Seq coverage ratio is plotted as the whole-fragment coverage ratio of micrococcal-nuclease-digested chromatin to randomly fragmented chromatin in sliding, 500-bp windows (step size, 1 bp).

different binding specificities of its alleles[25–27]. Approximately 40% of human recombination hotspots possess a consensus motif[28] that matches the predicted binding site of human PRDM9. We were able to identify a consensus motif specific to mouse hotspots (Fig. 2d, e). Sequences with better alignment scores to the motif consensus are more strongly over-represented in the hotspot regions, with the best hits showing almost 180-fold enrichment (Supplementary Fig. 10). Overall, hotspots containing consensus sequences are stronger than those without, and the quality of the motif alignment within hotspots is positively correlated with hotspot strength (Supplementary Fig. 11). Importantly, the motif shows a strong match to the predicted binding site of the *Prdm9* allele present in our mouse strain (Fig. 2d and Supplementary Fig. 12), and it is present at the centre of at least 73% of hotspots. This indicates that PRDM9 is a determinant for many more hotspots than previously thought and can explain the stronger-than-expected correlation between hotspot activity and *Prdm9* allelic variation found in recent association[27] and sperm typing[29] studies.

Because post-translational histone modification has been implicated in the regulation or maintenance of recombinational activity, we asked whether DNA in the hotspot regions is assembled into nucleosomes. Strikingly, we found that both the predicted[30] and the actual nucleosomal occupancies are co-centred with recombination hotspots (Fig. 2f), reflecting a previously unknown, intrinsic ability of hotspot DNA to assemble a nucleosome. We next examined the distribution of the H3K4me3 marks in germ (testis) and somatic (liver) tissues. We found that 94% of hotspots overlap H3K4me3, with the majority of hotspots overlapping testis-specific H3K4me3 marks (87%) and practically none overlapping the marks specific to the liver (Fig. 3a and Supplementary Fig. 13). Although the association of H3K4me3 with DSB hotspots is not surprising in light of recent work[23–27], the extent of the overlap revealed here is such that H3K4me3 can be considered a global feature of DSB sites in multicellular organisms. Importantly, unlike in *Saccharomyces cerevisiae*[23], where gene promoters and hotspots seem to share the same H3K4me3 mark, H3K4me3 marks at mouse hotspots are hotspot specific (Fig. 3b). Even when a hotspot is located very close to a transcription start site, the corresponding H3K4me3 marks are clearly spatially distinct (Fig. 3c), indicating that different mechanisms are involved in the histone modifications at these sites (Fig. 3c, Supplementary Fig. 13 and Supplementary Text).

Importantly, H3K4me3 per se is not a sufficient mark for DSB formation. Why meiotic DSBs are correlated with only a small fraction (16.7%) of the ~55,000 H3K4me3 marks in testis remains a mystery. An attractive possibility is that some component of the DSB machinery directly associates with PRDM9 or a PRDM9-containing complex and is therefore delivered to the potential DSB sites. Subsequent trimethylation of H3K4 might be required to set the stage for recombination initiation and progression. Other possibilities may also be considered: (i) in addition to H3K4me3, other epigenetic marks may be present that require H3K4me3 to allow DSB formation; (ii) H3K4me3 introduced

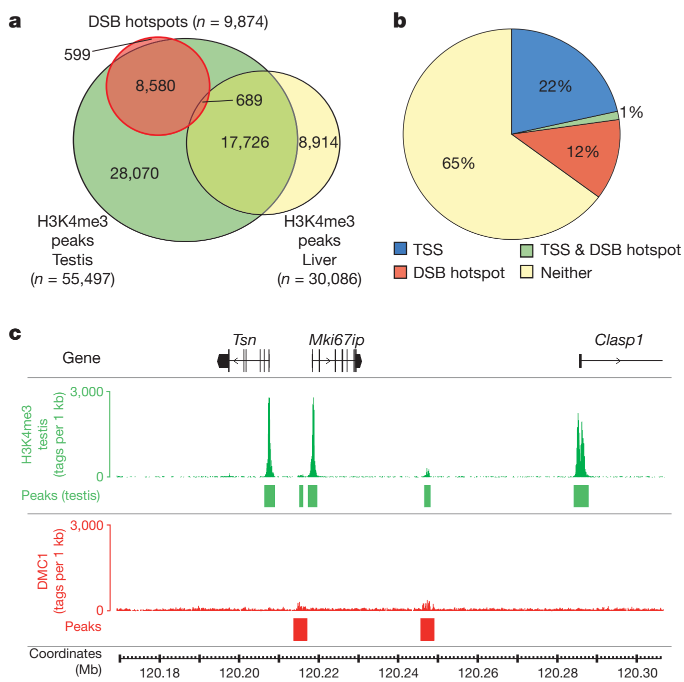


Figure 3 | Specific H3K4me3 marks are associated with DSB hotspots. **a**, The vast majority (93.9%) of DSB hotspots overlap H3K4me3 marks, most of which (86.9%) are testis specific. The six DSB hotspots that overlap liver-specific H3K4me3 marks are not shown. Peak calling for each data set was performed using an equal number of tags. **b**, DSB hotspots are associated with a set of H3K4me3 marks that are distinct from those at transcription start sites (TSSs). The fraction of H3K4me3 marks overlapping TSSs and/or DSB hotspots, or neither, is shown. **c**, H3K4me3 marks at DSB hotspots are generally weaker than TSS-associated marks, and are also spatially distinct despite being sometimes in very close proximity. Tag coverage is displayed in 100-bp steps.

by PRDM9 may be different from other H3K4me3 marks (for example, PRDM9, unlike other methyltransferases, might modify one rather than both H3 histones in the same nucleosome, or the other way around); (iii) specific histone variants could be present in the PRDM9-modified nucleosomes or could be substrates of PRDM9; (iv) transcription factors or other proteins bound to 'non-PRDM9' trimethylation marks may interfere with the DSB machinery.

METHODS SUMMARY

We used *Hop2*$^{-/-}$ mice[13] on a [C57BL/10.S $\times$ C57BL/10.F] F_1 genetic background and wild-type mice on the same background to make a map of DSB hotspots. ChIP and high-throughput sequencing were performed according to manufacturer-provided protocols (Upstate and Illumina, respectively) with minor modifications. For each sample, 36-bp end sequences were aligned to the mm9 reference genome using the Illumina GAII analysis pipeline. Only quality-filtered reads that mapped uniquely to the genome were retained for downstream analyses. We identified DSB hotspots and H3K4me3 peaks by comparing the sequence tag coverage for each ChIP sample with that of tag-count-matched control samples using the MACS algorithm. The hotspot consensus motif was identified from nonamers enriched near hotspot centres using a bespoke analysis pipeline. Additional details are available in Supplementary Information.

Received 5 October 2010; accepted 24 January 2011.
Published online 3 April 2011.

1. Arnheim, N., Calabrese, P. & Tiemann-Boege, I. Mammalian meiotic recombination hot spots. *Annu. Rev. Genet.* **41,** 369–399 (2007).
2. Buard, J. & de Massy, B. Playing hide and seek with mammalian meiotic crossover hotspots. *Trends Genet.* **23,** 301–309 (2007).
3. Lichten, M. Meiotic chromatin: the substrate for recombination initiation. *Genome Dynam. Stabil.* **3,** 165–193 (2008).
4. Paigen, K. & Petkov, P. Mammalian recombination hot spots: properties, control and evolution. *Nature Rev. Genet.* **11,** 221–233 (2010).
5. Clark, A. G., Wang, X. & Matise, T. Contrasting methods of quantifying fine structure of human recombination. *Annu. Rev. Genom. Hum. Genet.* **11,** 45–64 (2010).
6. Kauppi, L., May, C. A. & Jeffreys, A. J. Analysis of meiotic recombination products from human sperm. *Methods Mol. Biol.* **557,** 323–355 (2009).
7. The International HapMap Consortium. A haplotype map of the human genome. *Nature* **437,** 1299–1320 (2005).
8. Durbin, R. M. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467,** 1061–1073 (2010).
9. Frazer, K. A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449,** 851–861 (2007).
10. Kong, A. *et al.* Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467,** 1099–1103 (2010).
11. Myers, S., Bottolo, L., Freeman, C., McVean, G. & Donnelly, P. A fine-scale map of recombination rates and hotspots across the human genome. *Science* **310,** 321–324 (2005).
12. Neale, M. J. & Keeney, S. Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* **442,** 153–158 (2006).
13. Petukhova, G. V., Romanienko, P. J. & Camerini-Otero, R. D. The Hop2 protein has a direct role in promoting interhomolog interactions during mouse meiosis. *Dev. Cell* **5,** 927–936 (2003).
14. Qin, J., Richardson, L. L., Jasin, M., Handel, M. A. & Arnheim, N. Mouse strains with an active H2-Ea meiotic recombination hot spot exhibit increased levels of H2-Ea-specific DNA breaks in testicular germ cells. *Mol. Cell. Biol.* **24,** 1655–1666 (2004).
15. Paigen, K. *et al.* The recombinational anatomy of a mouse chromosome. *PLoS Genet.* **4,** e1000119 (2008).
16. Cox, A. *et al.* A new standard genetic map for the laboratory mouse. *Genetics* **182,** 1335–1344 (2009).
17. Burgoyne, P. S. Genetic homology and crossing over in the X and Y chromosomes of mammals. *Hum. Genet.* **61,** 85–90 (1982).
18. Khambata, S., Mody, J., Modzelewski, A., Heine, D. & Passmore, H. C. Ea recombinational hot spot in the mouse major histocompatibility complex maps to the fourth intron of the Ea gene. *Genome Res.* **6,** 195–201 (1996).
19. Francino, M. P. & Ochman, H. Strand asymmetries in DNA evolution. *Trends Genet.* **13,** 240–245 (1997).
20. Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genom. Hum. Genet.* **10,** 285–311 (2009).
21. Mihola, O., Trachtulec, Z., Vlcek, C., Schimenti, J. C. & Forejt, J. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* **323,** 373–375 (2009).
22. Wang, Z., Schones, D. E. & Zhao, K. Characterization of human epigenomes. *Curr. Opin. Genet. Dev.* **19,** 127–134 (2009).
23. Borde, V. *et al.* Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO J.* **28,** 99–111 (2008).
24. Buard, J., Barthes, P., Grey, C. & de Massy, B. Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. *EMBO J.* **28,** 2616–2624 (2009).
25. Parvanov, E. D., Petkov, P. M. & Paigen, K. *Prdm9* controls activation of mammalian recombination hotspots. *Science* **327,** 835 (2010).
26. Myers, S. *et al.* Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* **327,** 876–879 (2010).
27. Baudat, F. *et al.* PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* **327,** 836–840 (2010).
28. Myers, S., Freeman, C., Auton, A., Donnelly, P. & McVean, G. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nature Genet.* **40,** 1124–1129 (2008).
29. Berg, I. L. *et al. PRDM9* variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nature Genet.* **42,** 859–863 (2010).
30. Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458,** 362–366 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Lichten (NCI, NIH) and P. Hsieh (NIDDK, NIH) for comments and discussion. We are grateful to S. Sharmeen for her help with high-throughput sequencing. This work was supported in part by Basil O'Connor Starter Scholar Research Award Grant No. 5-FY07-667 from the March of Dimes Foundation (G.V.P.); NIH grant 1R01GM084104-01A1 from NIGMS (G.V.P.); New Investigator Start-up Grants FS71HU, R071HU and CS71HU from USUHS (G.V.P.); and the NIDDK (NIH) Intramural Research Program (R.D.C.-O.).

Author Contributions F.S. performed all experiments. I.V.G., K.B. and P.K. performed computational data analyses. All authors contributed to experimental design. G.V.P. and R.D.C.-O. designed and supervised the study. G.V.P. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information ChIP-Seq data have been deposited in the Gene Expression Omnibus under accession number GSE24438. *Prdm9* complementary DNA sequences have been deposited in GenBank under accession numbers HQ704390 and HQ704391. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.V.P. (gpetukhova@usuhs.mil) or R.D.C.-O. (rdcamerini@mail.nih.gov).

COLUMN PhD scientists can learn from the woes of law and medical students **p.381**

CAREERS BLOG The latest discussions and news on research jobs **go.nature.com/z8g4a7**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

A. RUGGIERI/ILLUSTRATION WORKS/CORBIS

START-UPS

In search of venture capital

To secure elusive funding, entrepreneurs must understand the financial landscape and the motivations of investing firms.

BY SARAH KELLOGG

Somewhere between borrowing money from friends and family and maxing out their personal credit cards, many scientists and researchers trying to take their discoveries from the lab to the marketplace decide to seek the counsel and financial support of a venture-capital firm. Such firms have long had a key role in establishing legal structures and marketing strategies for start-up companies, as well as in providing them with the funding to stay afloat until they are robust enough to secure a commercial bank loan, be purchased by a larger competitor or achieve sales that signal long-term success.

Yet, like insurance-claims adjusters, venture-capital firms are often the kind of friends that scientists might not want. That doesn't mean that they aren't highly valued partners, but the relationship frequently starts out on an unequal and problematic footing. "It's not just about money, it's about chemistry," says Ellen Rudnick, executive director of the Michael P. Polsky Center for Entrepreneurship at the University of Chicago in Illinois. "You want somebody there whom you trust and who really understands your business. You have to think of them as a marriage partner."

Scientists can improve their odds of success in securing capital by knowing their options and responsibilities. The likelihood of achieving victory and avoiding pitfalls, such as surrendering one's fledgling company, increases if researchers know their goals from the outset, put a plan into motion to achieve them, recruit established hands to support their pursuit and go into the search with realistic expectations.

DEVISING A GAME PLAN

Firms aren't drawn to great ideas alone; they're attracted to great ideas that have the promise of financial success. It's a painful truth that many astonishing innovations won't ever win the favour of venture-capital firms, because they lack a market. "You need to ask yourself from a purely scientific standpoint: is this idea differentiated in the eyes of my peers? Is there something unique scientifically about my idea?" says Robert Nelsen, co-founder and managing director of ARCH Venture Partners in Seattle, Washington. "We're looking for the revolutionary rather than the evolutionary."

For scientists working in universities or government labs, the vetting process starts with technology-transfer offices, which protect, manage and license research. Scientists who receive university or federal grants need to work out any potential licensing sticking points before embarking on a venture-capital search. Firms want to ensure that scientists are free to develop the discovery, and will need to determine whether the licensing of the technology is exclusive to one company, and for how long. When working with university-based scientists, the firm will also want to know whether the institute receives equity in the new company and whether the licence reverts back to the university if the start-up does not meet certain milestones or financial requirements.

A good technology-transfer office will employ business experts with experience in developing pitches and building management teams for start-ups. They can help in protecting discoveries, filing patent applications ▶

and outlining and resolving the researcher's licensing concerns.

But the challenges don't end there. Eva Harth, a chemist at Vanderbilt University in Nashville, Tennessee, who is trying to market a degradable nanosponge for the treatment of cancer and eye diseases, says that her search for the right firm to optimize her invention's commercial potential has been a struggle. "Getting from the science to putting it into the hands of physicians has been challenging, especially since I'm not an expert in business or finance," says Harth. "If you don't have the connections and don't know how the system works, you're just lost."

VENTURE COLLAPSE

To get a handle on the current financial landscape, scientists need to understand the impact of the global economic crisis. The financial collapse of 2008 not only shrank the size and availability of investment capital, but also reduced the number of venture-capital firms. In the United States alone, the number dropped from 996 in 2007 to about 791 in 2010, and the threat of closures continues.

The largest impact on venture-backed companies may have come from the declining initial-public-offering market, says Emily Mendell, vice-president of communications for the National Venture Capital Association (NVCA) in Arlington, Virginia, the trade group for US venture-capital firms. Companies that had been in a position to go public were forced to wait, says Mendell, so firms had to invest more money and time than they had planned to keep promising companies afloat while the market righted itself. "There will be firms that will not be able to raise follow-on funds," says Mendell. "We are seeing a shrinking industry going forward." She notes, however, that the venture-capital industry in the United States, although smaller than it was, is not moribund. That was demonstrated in 2010, which saw the first year-on-year rise in investments since 2007. Venture-capital firms invested US$21.8 billion in 3,277 deals last year: a 19% boost in dollars compared with 2009 and a 12% increase in the number of deals, according to the 2010 *MoneyTree Report* by the NVCA and PricewaterhouseCoopers, a financial-services firm based in London.

Still, the current status of the industry helps to explain why many scientists feel that firms are increasingly reluctant to fund projects without a guaranteed return on their investment. "Venture implies that they take risk, but I think many of these firms take no risk at all," says Harth. "They are looking for guarantees only."

The United States and Europe are important centres for innovation and investment, and look set to remain so. But emerging markets are expected to drive the industry over the next five years, according to the *2010 Global Venture Capital Survey* by the NVCA and Deloitte of New York. The survey, which measures the opinions of more than 500 venture capitalists worldwide, found that 92% in the United States think that the number of venture-capital firms in the country will decline in the next five years, a view echoed in France, Israel and the United Kingdom. Meanwhile, 99% of venture capitalists in China, 97% in Brazil and 85% in India expected to see an increase in numbers of local venture-capital firms.

The upshot: scientists in less-developed countries will probably see rising investment opportunities through private entities. But there are some caveats attached to the boom. Venture capitalists in emerging markets may be inexperienced — leaving scientists ceding control to firms not familiar with regulations — or ill-prepared to raise the large amounts of capital required for biotechnology and other health-related investments. With so much at stake, it is essential that scientists exercise due diligence in selecting a venture-capital firm.

KNOCKING ON DOORS

Before selection begins, experts say, scientists should avoid trading away pieces of their company to less-experienced individual entrepreneurs and angel-fund investors for early investment funds before approaching venture-capital firms. If scientists have already signed binding agreements with other funders or licence holders, firms will see only unappealing prospects: limited profitability and months of legal negotiations. "The biggest mistake faculty members make is to partner with entrepreneurs who are not of the quality or experience that venture investors will accept," says Nelsen. And it is often best to secure only the investment needed for the next 12–18 months, rather than seek funding for an extended period of time. To win more-substantial, long-term funding for a smaller, less-established company often means trading away more of the enterprise than it would once the company has recruited its first client or validated its first innovation.

Once that concern is dodged, the capital–start-up relationship must be considered. If it's akin to marriage, then the first step in establishing it is finding and courting a firm that fits the personality of the technology or discovery at the heart of the fledgling company. By studying a firm's track record, scientists can learn about the type of company that they support and the successes and failures they've had in the past. It helps to call a few colleagues or university entrepreneurship officers to gauge a firm's reputation among scientists in the field, because word travels fast about flawed and failing venture-capital firms.

More importantly, researchers should look for that vital link, person or connection that can open the door. Firms say that for every 100 proposals they receive from researchers and entrepreneurs, they will fund only a single start-up. Being taken seriously as a prospect often starts with a good reference from a colleague or department chair who has the credentials to attract high-powered investment. "In any university or national lab, find the person who has had success in creating a start-up, and they most likely can introduce you to someone," says Nick Galakatos, managing director of Clarus Ventures in Cambridge, Massachusetts.

Once they have a foot in the door, researchers need to make a compelling case. They must explain what need their product fulfils, and how they intend to market and sell it. These preliminary conversations, which must be both aspirational and grounded in reality, help the firm to understand the level of risk involved in the prospective investment. They also provide early opportunities to assess the business acumen of the scientists on the project.

"It's not just about money, it's about chemistry."
Ellen Rudnick

The best firms treat entrepreneurs as important customers and add tremendous value to a start-up in terms of recruiting, strategy, coaching and connections. But they are not doing so out of the goodness of their hearts, says serial entrepreneur Steve Blank, a lecturer in entrepreneurship at the University of California, Berkeley, and author of a blog for start-ups (http://steveblank.com). "Entrepreneurs need to understand that VCs are simply a sophisticated form of financial investors who in turn need to satisfy their own investors," he wrote on his blog this year. Some investors acknowledge that there are bad actors in the field, who mistreat scientists and want to wrest companies away from their founders. But scientists' general lack of business experience makes them ill-suited to run multi-million-dollar companies in competitive environments for themselves. Furthermore, any firm willing to invest heavily in a start-up will expect to guide and manage that company's future.

Ultimately, scientists should take heart in the market forces that drive venture-capital firms. Those firms that fail to invest robustly in science, or that treat scientists unfairly, only damage their own reputations and undermine their own success. "Our business is to make people money, and that includes entrepreneurs and scientists," says Nelsen. "If we make a huge company that is quite successful and the scientist doesn't make money, and the university doesn't make money, that's a huge failure for us. We want people to keep coming back to us again and again with their discoveries — so we all make money." ■

Sarah Kellogg *is a freelance writer in Washington DC.*

COLUMN

What is a PhD really worth?

An advanced degree doesn't always bring the prospects it once did, says Peter Fiske. But scientists can learn from the travails of those with professional qualifications.

In the past few months there has been a string of articles[1,2] in the Western press about the poor career prospects of graduates from professional schools, especially law school. Thanks to a dismal job market, applications to such institutions surged during the recession. Now, a few years later, those applicants are graduating into a career market that is only slightly better — and many have become saddled with debt.

Professional schools are now encountering the graduate-overproduction issues with which PhD programmes have been wrestling for decades. Entry to law school or medical school once provided a near-certainty of gainful (and often highly lucrative) employment after graduation, but the alumni of these programmes are now struggling to find their first jobs. Many are blaming their graduate schools for promulgating misleading information about their job prospects.

IMAGES.COM/CORBIS

Sound familiar? In the late 1980s, the US National Science Foundation and other groups spread the notion that a wave of retirement was about to sweep through academia, and that the academic job prospects for emerging PhDs had never been brighter. In fact, the economic assumptions that formed the basis of this prediction were erroneous[3], and no such wave of retirement took place. In the early 1990s, PhD holders (including myself) spilled into the job market with essentially no guidance on how to find employment outside the ivory tower. The ensuing backlash from young scientists forced the academic establishment to acknowledge that it was producing many more scientific PhDs than it could accommodate in tenured positions. More importantly, the establishment had to recognize that PhD scientists could find satisfying and valuable careers outside academia. Although many graduate-student scientists still have their hopes fixed on academic positions, they are now also exploring a range of career options in industry, government and the non-profit sector.

LESSONS FOR LIFE

The predicament of the current crop of law- and medical-school graduates might give PhD students an opportunity for *Schadenfreude*. But I think that young scientists — and the academic institutions from which they come — can also learn from the dilemma faced by budding lawyers and physicians.

The first lesson is that quality matters. Most of the lamentation from law-school graduates is coming from those who attended lower-ranked programmes. Although there is ample evidence that the current mechanisms for ranking graduate programmes are limited and problematic, there is nevertheless a high correlation between the overall ranking of a programme and the employment outcomes of its alumni. Therefore, prospective students should shop carefully and ask for real data on the career paths of each programme's graduates. In the United States, tools such as the *Grad School Guide* (http://graduate-school.phds.org) cast light on the actual quality and outcomes of specific graduate courses. And the US National Academy of Sciences has just released its latest ranking of doctoral programmes (http://sites.nationalacademies.org/pga/resdoc/index.htm).

The second lesson is that debt is important — and this is where the situation for science graduates deviates from that of medical and law students. Although graduate students in the sciences essentially sign themselves up for years of penury in university, nearly all US PhD programmes, and many others around the world, provide stipends for their researchers. PhD students may graduate with few possessions other than a couple of boxes of books and some used Ikea furniture, but at least they do not end up with a huge bill for years of tuition, as law and medical students often do. Frugal science graduate students emerge with a very important life lesson: money does not buy happiness (although it certainly can make misery a lot more comfortable), and living frugally is better than amassing tens of thousands of dollars of debt. Years spent with limited income teach PhD-holding scientists how to be resourceful — and that can be of enormous value throughout a career.

But I believe the most important lesson is that no programme of higher education can guarantee its graduates gainful and lucrative employment. At best, a graduate programme in any discipline can provide its students with key skills, knowledge and abilities. How the graduates apply that learning is up to them.

BROADENED HORIZONS

That being said, broader training would help graduates' chances. PhD programmes in the sciences still overemphasize the academic track and actively devalue other career paths. Expanding the PhD experience and preparing holders of scientific doctorates to be successful in a range of careers would not require a major overhaul of graduate programmes. Focused seminars in areas such as communication, business basics and public policy would go a long way towards strengthening the capabilities of PhD students and improving their career prospects. Few academic programmes fully appreciate the true potential that PhD training can confer, or the breadth and depth of value that someone with a PhD can contribute to the world at large; universities often believe that academia is still the most valuable calling for their graduates.

Graduate students and postdocs are in the best position to organize themselves and insist on such professional education as part of their training. After all, they stand to gain the most. ■ SEE NEWS FEATURE P.280

Peter Fiske *is chief executive of PAX Water Technologies in Richmond, California, and author of* Put Your Science to Work *(AGU, 2001).*

1. Segal, D. 'Is Law School a Losing Game?' *The New York Times* (8 January 2011). Available at http://go.nature.com/cf67wp
2. The disposable academic: why doing a PhD is often a waste of time. *The Economist* (6 December 2010). Available at http://go.nature.com/qu7c8z
3. Shapiro, D. T. *J. Higher Edu.* **72,** 532–564 (2001).

TEST OF FAITH

The heat is on.

BY BRIAN P. FRANK

I must not fail. Sixth-year graduate student Vikram Singh was sweating profusely, not just from the heat of the nearby bonfire or the glowing bed of coals spread before him. Around him, the faculty members and students of the Chemical Engineering Department clapped and chanted the ritual songs while Professor Markson leapt about in a frenzy as she led the undergraduates, her long grey hair released from its customary tight bun. Her face was streaked with ceremonial paint in red and black, the school colours. The old bat looked utterly terrifying, worse than she had during his candidacy exams. Vikram stood barefoot and wore only the traditional loincloth.

All I have to do is walk across the coals. My ancestors in India did this for hundreds of years simply through the power of their belief. So why not me?

Of course, most of India was under water now, and he had grown up a Greenhouse refugee in Kansas. But he had climbed out of the camps, found scholarships to get himself here to the Institute and onward for graduate work until only one final barrier remained. But if he couldn't pass this test, it would all be for nothing. The noise grew louder, Professor Markson leaping higher until Vikram was sure she was going to give herself a coronary. It suddenly stopped as Professor Li, Vikram's thesis adviser, strode into the firelight and began the ritual speech:

"We gather here tonight to test the faith of Vikram Singh. Since the time of George the Second" — there was a chorus of enthusiastic boos and hisses from the crowd — "we have known that it is not enough for scientists and engineers to choose cold detachment. Those were dark days of research suppressed, of scientists persecuted for speaking truths that could have saved us all. And what is the result? Global disaster so vast that we still live with the consequences. It must not happen again."

"Never again," the crowd murmured.

"We must stand against those who would promote ignorance for selfish ends. We must have true faith! Vikram Singh!" Li barked, "Where do you place your faith?"

Vikram licked dry lips and gave the response. "I place my faith in the facts as revealed by science."

Li went on: "Like the martyrs of science — Hypatia, Archimedes, Galileo, DeMorgan — we must be willing to give our very lives for our faith. Now you too must face this ultimate trial."

No one spoke. No one moved. This was their department's most sacred tradition, and everyone knew it. Other disciplines had their own rituals, but this was the Chemical Engineering way.

Then Li shouted: "Let the final test begin!"

JACEY

Vikram turned to face the coals as the crowd roared its approval. He'd told no one about that horrifying and unforgettable night when he was seven and the fire had raced unchecked through the camp.

If I believe strongly enough, then the coals won't burn me. He took a deep breath and without thinking stepped onto the bed of coals and began to walk across. *Just keep moving…*

He made it halfway across before the terror pounced. *I'm standing on a bed of red-hot coals!*

It was like that awful night in the camp, only a thousand times worse. The fear raged about him like a storm, and in his panic everything seemed to slow and stop around him. It was only a second or two of suspended time, but within it Vikram thought harder and faster than he ever had in his life.

He had to have faith; if he only had faith he could do this. *Where do you place your faith?* he asked himself, and automatically his mind answered, *I place my faith in the facts as revealed by science.* And then it hit him.

I believe in thermodynamics.

Thermal effusivity. Unsteady state heat flow. Specific heat capacities. Leidenfrost effect. The equations blazed right there in his mind, basics from a course drilled into every first-year student at the Institute. *The ash surrounding the coals has a low thermal conductivity so it acts as an insulator. The water in my tissues has a very high heat capacity but the coals have a very low one, therefore my foot's temperature will change less than that of the coals. Water has a high thermal conductivity so the heat will be carried away from the bottom of my feet; coals have a low thermal conductivity so the only heat comes from the surface of the coals actually in contact with my feet.* You didn't need faith to believe that firewalking worked. All you really needed was to understand thermodynamics.

No, it was more than that. This wasn't just about his understanding, it really was about his faith. Not his faith in his beliefs; that was the mistake they had made in the last century, that somehow their beliefs about the facts would change them. Instead, this was a test of his belief in science and its ability to reveal the facts. So did he truly believe in the facts as revealed by science? Enough to walk across fire?

The moment of paralysis stretched and broke, but he knew his answer now and confidence surged through him. Unafraid, he crossed the remainder of the coals to collapse at last, smiling and unhurt, into the arms of the crowd. Then Li pushed through and grasped Vikram's hands and held them high over his head like a championship boxer as he shouted: "I present to you Vikram A. Singh, graduate student no more. Congratulations, Vikram A. Singh, PhD!" ■

➲ NATURE.COM
Follow Futures on Facebook at:
go.nature.com/mtoodm

Brian P. Frank *holds a PhD in chemical engineering, and conducts research into aerosol-phase environmental nanoparticles. This story is dedicated to his late mentor, thermodynamics professor Hank Van Ness.*